4-1-2013

Characterizing Financial and Statistical Literacy

Amalia Girolamo
Durham University Business School

Glenn Harrison
Georgia State University

Morten Lau
Durham University Business School

Todd Swarthout
Georgia State University

Characterizing Financial and Statistical Literacy

by

Amalia Di Girolamo, Glenn W. Harrison, Morten I. Lau and J. Todd Swarthout †

April 2013

ABSTRACT.

We characterize the literacy of an individual in a domain by their elicited subjective belief distribution over the possible responses to a question posed in that domain. We consider literacy across several financial, economic and statistical domains. We find considerable demographic heterogeneity in the degree of literacy. We also characterize the degree of consistency within a sample about their knowledge, even when that knowledge is imperfect. We show how uncertainty aversion might be a normatively attractive behavior for individuals who have imperfect literacy. Finally, we discuss extensions of our approach to characterize financial capability, the consequences of non-literacy, social literacy, and the information content of hypothetical survey measures of literacy.

† Department of Accounting and Finance, Durham University Business School, UK (Di Girolamo and Lau); Department of Risk Management & Insurance and Center for the Economic Analysis of Risk, Robinson College of Business, Georgia State University, USA (Harrison); Department of Economics, Andrew Young School of Policy Studies, Georgia State University, USA (Swarthout). E-mail contacts: amalia.digirolamo@durham.ac.uk, gharrison@gsu.edu, m.i.lau@durham.ac.uk and swarthout@gsu.edu. We are grateful to the Danish Social Science Research Council (grant #12-130950) and the Center for Actuarial Excellence Research Fund of the Society of Actuaries for financial support.

Table of Contents

When we say that someone is literate we mean more than that they can just "read and write." The term more generally indicates someone who is educated, whether by formal or informal means, and able to comprehend topics through words.[1] Characterizing and measuring the literacy of an individual requires then that we have some way of assessing how knowledgeable the person is about certain topics. There are some topics about which one can have "crisp" knowledge, in the sense of Boolean truth values. However, there are many domains of knowledge that one naturally expects varying levels of precision. We characterize literacy in terms of the subjective beliefs that someone has over possible responses to some question.

Following Savage [1971][1972], we *define* subjective beliefs by the choices that individuals make when facing bets whose outcomes depend on those beliefs. The measurement of the literacy that someone has in a specific domain entails the elicitation of their subjective beliefs. For that task we will conduct an experiment using proper scoring rules, which are simply structured bets offered to the individual by an observer (the experimenter). All of the elicited beliefs were incentivized and incentive-compatible, so that the subjects were making real choices with real economic consequences.

Our approach is to elicit the entire subjective belief *distribution* that an individual has, to ascertain how precise their knowledge is in response to some question. This generalizes the prevailing approach to measuring literacy, which considers responses to (hypothetical) multiple choice questions (e.g., Lusardi and Mitchell [2007][2008]). For a specific question or domain, we are able to say "how literate" the person is, rather than just say that they are or are not literate. Of course, by asking a series of questions one can ascertain the fraction of correct answers for an individual with the traditional approach, but that requires one to pool responses over different questions which may span different

[1] The *Oxford English Dictionary (Second Edition)* defines the adjective "literate" as someone who is "acquainted with letters or literature; educated, instructed, learned." Remund [2010] offers a balanced account of the many definitions of literacy found in the academic and policy literature. Our focus on financial knowledge corresponds to the first of his categories of conceptual definitions of literacy (p. 279).

knowledge domains.

The domains of interest to us are financial, economic and statistical knowledge. We consider a mixture of questions in which the correct answer involves the application of logical and grammatical principles, and questions in which the correct answer involves some specific fact. This reflects a trend in the measurement of literacy towards more than just the ability to draw logical or grammatical conclusions from information presented in the question itself, and to also consider awareness of facts that are of importance for the functioning of the individual.

A byproduct of our characterization is that we can also say something about the degree of common knowledge that a sample of individuals have about some proposition. Quite apart from whether or not a given individual knows the true answer with some precision, we often want to know if a group of individuals have the same degree of knowledge. In effect, we are able to operationalize several interpretations of what it means to have heterogeneous beliefs.

In section 1 we describe the experimental task we developed and employed with 120 subjects. In section 2 we review in detail the properties of the subjective belief elicitation procedure. In section 3 we present results on the degree of financial and statistical literacy of our subjects. In section 4 we consider the consistency of knowledge across subjects. Section 5 considers important extensions of our approach, and section 6 concludes.

1. Procedures

A. Literacy

We consider literacy in terms of 8 specific questions asked of each subject in an experiment. In each case there is a correct answer, and responses were elicited over a continuous range of possible answers presented in terms of 10 intervals or "bins." A computer interface was used to present the belief elicitation tasks to subjects and record their choices, allowing them to allocate tokens in accordance with their subjective beliefs. Figure 1 presents the interface.[2] The interface implements the Quadratic Scoring Rule discussed in section 2. Subjects could move the sliders at the bottom of the screen to re-allocate the 100 tokens as they wished, ending up with some distribution. The instructions explained that they could earn up to £20, as shown in Figure 2, but only by allocating all 100 tokens to one interval *and* that interval containing the true answer: if the true answer was just outside the selected interval, they would in that case receive £0. At the time of the experiments in December 2012, £20 was worth roughly $32.

Subjects were rewarded for one of these 8 belief elicitation tasks, with the task selected at random. The correct answer was revealed, and their earnings calculated according to the number of tokens allocated to the true interval in their elicited beliefs. For example, if the respondent had reported the beliefs in Figure 1, she would have been paid £16.25 if the correct answer was between 8% and 9.99%. As it happens, the correct answer here is 7.9%, so the subject would have actually received £11.25 since the correct answer was in the next lower interval, corresponding to unemployment rates between 6% and 7.99%.

The incentivized questions were as follows:

- **Q1: Interest Compounding**. "Suppose you had £100 in a savings account and the interest rate is 2% per year and you never withdraw money or interest payments. After 5 years, how much would you have on this account in total?" The correct answer is £110.40, and responses were elicited between £105 and £115 in intervals of £1.
- **Q2: Inflation**. "Suppose you had £200 in a saving account. The interest rate on your saving account was 1% per year and inflation was 2% per year. After 1 year, how much would be the value of the money on this account?" The correct answer is £198, and responses were elicited between £195 and £205 in intervals of £1.
- **Q3: Expected Lifetime for Men**. "Based on 2010 National Statistics, if a man lived to be 20 in the United Kingdom, how many more years would he expect to live? Note that this is not the age he would die at, but how many more years he would expect to live." The correct answer is 59.1 years, and responses were elicited between 0 and 100 years in intervals of 10

[2] The instructions are reproduced in full in Appendix A. The interface was initialized with 10 tokens allocated to each bin.

years.[3]

- **Q4: Expected Lifetime for Women**. "Based on 2010 National Statistics, if a woman lived to be 20 in the United Kingdom, how many more years would she expect to live? Note that this is not the age she would die at, but how many more years she would expect to live." The correct answer is 62.9 years, and responses were elicited between 0 and 100 years in intervals of 10 years.
- **Q5: Breast Cancer Risk Presented in Probability Format**. "The probability of breast cancer is 1% for women at age 40 who participate in routine screening. If a woman has breast cancer, the probability is 80% that she will get a positive mammography. If a woman does not have breast cancer, the probability is 9.6% that she will also get a positive mammography. A woman in this age group had a positive mammography in a routine screening. What is the probability that she has breast cancer?" The correct answer is 7.8%, and responses were elicited between 0% and 100% in intervals of 10%.
- **Q6: Breast Cancer Risk Presented in Frequency Format**. "10 out of every 1000 women at age 40 who participate in routine screening have breast cancer. 8 of every 10 women with breast cancer will get positive mammography. 95 out of every 990 women without breast cancer will also get a positive mammography. In a new representative sample of 100 women at age 40 who got a positive mammography in routine screening, how many of these women do you expect to have breast cancer?" The correct answer is 7.8, and responses were elicited between 0 and 100 in intervals of 10.
- **Q7: Credit Solvency Risk Presented in Probability Format**. "The probability of being technically insolvent is 10% for homeowners who participate in a credit score test. If a homeowner is insolvent, the probability is 90% that he will get a negative credit score. If a homeowner is solvent, the probability is 5% that he will get a negative score. A homeowner had a positive credit score in a routine test. What is the probability that he is insolvent?" The correct answer is 66.6%, and responses were elicited between 0% and 100% in intervals of 10%.
- **Q8: Credit Solvency Risk Presented in Frequency Format**. "100 out of every 1000 homeowners who participate in routine credit score tests are technically insolvent. 90 of every 100 insolvent homeowners will get negative credit scores. 45 out of every 900 solvent homeowners will also get a negative credit score. In a new representative sample of 100 homeowners who got a negative credit score in routine tests, how many of these homeowners do you expect to be insolvent?" The correct answer is 67, and responses were elicited between 0 and 100 in intervals of 10.

The order of presentation of questions was randomized for each subject.

The first two questions are natural extensions of questions asked by Lusardi and Mitchell

[3] At the time of the experiment we did not have access to the correct Life Tables, and instead subtracted 20 from the expected lifetime at birth from the UK Office of National Statistics to pay subjects that had this question selected. The difference for aggregates such as "all men" or "all women" is tiny, and did not affect the payments to any subject given that the bin intervals we used were in 10-year increments: there is a difference of 0.6 of a year for men (59.1 versus 58.5), and 0.5 of a year for women (62.9 versus 62.4).

[2007][2008] in the *Health & Retirement Survey* (HRS) of 2004 in the United States.[4] This survey is naturally representative of Americans over the age of 50. Our Q1 adapts the following question of theirs: "Suppose you had $100 in a savings account and the interest rate was 2 percent per year. After 5 years, how much do you think you would have in the account if you left the money to grow: more than $102, exactly $102, less than $102?" The main difference is that we ask for beliefs about the true answer over a wide range. Our Q2 adapts this question of theirs: "Imagine that the interest rate on your savings account was 1 percent per year and inflation was 2 percent per year. After 1 year, would you be able to buy more than, exactly the same as, or less than today with the money in this account?" Lusardi and Mitchell [2012; Table 2.1] report that only 67.1% and 75.2% of their sample gave the correct response to each question, respectively. These fractions drop significantly (their Figures 2.1a and 2.1b) as one considers Black and Hispanic respondents. When the same questions were posed to a nationally representative sample of young Americans, aged between 22 and 28 in Wave 11 of the *National Longitudinal Survey of Youth* conducted in 2007-2008, 79.3% and 54.0% gave the correct responses to the interest rate and inflation questions, respectively (Lusardi, Mitchell and Curto [2010; Table 1, p. 365]).[5]

The next two questions ask about a basic informational input to retirement planning: expected remaining lifetime, conditional on reaching the age of 20. Indeed, Smith, Taylor and Sloan [2001; p. 1126] call this "the most important subjective risk assessment a person can make," although they were referring to own-mortality. We separate out the question for men and women, to ascertain if the

[4] A third question they asked was: *Do you think that the following statement is true or false? "Buying a single company stock usually provides a safer return than a stock mutual fund."* This question was posed in order to understand if the individuals know how to diversify their investment. In a later Dutch national survey van Rooij, Lusardi and Alessie [2011] increased the set of questions posed to individuals. Apart from 5 questions aimed at characterizing "basic" financial literacy (p. 452), they added 11 questions to characterize "advanced" financial literacy (p. 454). Similar extensions were undertaken by Bateman, Eckert, Geweke, Louviere, Thorp and Satchell [2012] in surveys in Australia.

[5] Bateman, Eckert, Geweke, Louviere, Thorp and Satchell [2012] ask these questions of adult retirement savers in Australia, and find that 78.4% get the inflation question correct and 71.8% get the interest rate question correct.

differential expected mortality between the two is recognized by individuals. These questions do not condition on the health, income, or any other relevant characteristics of the individual that would affect expected mortality. One could easily extend these questions to elicit more precise beliefs about someone who is more similar to the subject.

The most widely used subjective beliefs about longevity come from the *Health and Retirement Survey*, which has asked a simple question since 1992 to respondents under the age of 65: "With 0 representing absolutely no chance, and 100 absolute certainty, what is the chance that you will live to be 75 years of age or older?" A comparable question asks the chance that they would live to be 85, and for respondents over 65 a variant asked the chances of them living 11-15 years more. In the 2006 wave of the *Health and Retirement Survey* a sub-sample was asked questions that elicited their beliefs about the population life tables: "Out of a group of [men/women] your age, how many do you think will survive to the age of X?" The value of X was 75 for those under 65 themself, and 11-15 years older for those over 65. These questions are closer to those we asked, although we only conditioned on the single age 20.

Of course, these questions are not incentivized, and do not elicit information on the confidence of the subjective belief. However, Smith, Taylor and Sloan [2001] show that responses to this question are reasonably good predictors of future, actual mortality, even if they do not perfectly reflect new health information when updated. Perozek [2008] makes an even stronger case for the predictive value of these subjective belief questions, arguing that responses to these questions actually outperform population life tables. In contrast, Elder [2013] stresses that only with the 2006 wave can one evaluate the actual predictions, as early respondents reach the target ages of 75 or 85. And in that respect he presents a sharply contrary view, arguing that the evidence supports a "flatness bias," a "tendency for individuals to understate the likelihood of living to relatively young ages while overstating the likelihood of loving to ages beyond 80." He attributes this bias to a failure to recognize that mortality risk increases with age.

The last four questions present Bayesian updating questions to evaluate that dimension of statistical literacy. The first pair of questions on the risk of breast cancer are taken from the 15 questions considered by Gigerenzer and Hoffrage [1995; Table 2, p. 693]. Our Q5 and Q6 are direct translations of the parameters they use for the same breast cancer risk question. The difference between Q5 and Q6 is that the former uses conventional probability information to set up the problem, and the latter uses natural frequencies to set up the same problem.[6] Gigerenzer and Hoffrage [1995; Figure 3, p. 694] report a dramatic improvement in the correct application of Bayes Rule as their subjects moved from the probability format to the frequency format. For the breast cancer risk example, the fraction increases from about 15% to about 40%. Over all of their 15 problems, the fraction increased dramatically from 16% to 46%.

Our questions Q7 and Q8 consider another application of Bayes Rule reasoning, but with a financial example having to do with the risk of consumer credit solvency. This issue is particularly important in the United Kingdom, which had relatively high levels of consumer debt coming into the global recession of 2008. It has also been an active area of policy debate, as illustrated by the *Consumer Credit and Personal Insolvency Review* of the U.K. Treasury in November 2011.[7]

B. Demographics and Additional Measures

Apart from these incentivized subjective belief questions, which are our main focus, we asked subjects several hypothetical questions that have been widely used in the literature on cognitive abilities. One is **Cognitive Reflection Test** (CRT) proposed by Frederick [2005], consisting of three

[6] We use what they refer to as the Standard Probability format and the Standard Frequency format. They also considered shorter versions of each. We only report their results for the versions comparable to ours.

[7] See http://www.hm-treasury.gov.uk/fin_reform_consumer.htm.

questions:

- A bat and a ball cost £1.10 in total. The bat costs £1.00 more than the ball. How much does the ball cost?
- If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?
- In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?

The correct answers are £0.05, 5 minutes and 47 days. Many subjects responding to these hypothetical questions fail to "reflect" on some aspect of the information provided.[8] The other hypothetical battery is known as the **Berlin Numeracy Test**, and is due to Cokely, Galesic, Schulz, Ghazal and Garcia-Retamero [2012]:

- Imagine we are throwing a five-sided die 50 times. On average, out of these 50 throws how many times would this five-sided die show an odd number (1, 3 or 5)?
- Out of 1,000 people in a small town 500 are members of a choir. Out of these 500 members in the choir 100 are men. Out of the 500 inhabitants that are not in the choir 300 are men. What is the probability that a randomly drawn man is a member of the choir? (Please indicate the probability in percentage).
- Imagine we are throwing a loaded die (6 sides) 70 times. The probability that the die shows a 6 is twice as high as the probability of each of the other numbers. On average, out of these 70 throws, how many times would the die show the number 6?
- In a forest 20% of mushrooms are red, 50% brown and 30% white. A red mushroom is poisonous with a probability of 20%. A mushroom that is not red is poisonous with a probability of 5%. What is the probability that a poisonous mushroom in the forest is red?

The correct answers are 30, 25%, 20 and 50%.

In addition, a standard list of demographic questions were posed. These included age, sex, racial group, field of study, year of study, highest level of formal education expected to complete, current grade, citizenship, marital status, number of people in household, total income of the household, total income of parents, and smoking status.

---

[8] It is an interesting question whether one should care if subjects get the CRT questions right when they face incentives. In fact, many subjects do. But it could be argued that the CRT is designed to detect cognitive propensities to spot "minimally hidden" information in decision settings, and that this propensity is better detected when higher-order, executive brain functions are not engaged to earn money.

C. Risk Attitudes

We were interested in measuring risk attitudes for two reasons. The first was to ensure that the sample indeed was only modestly risk averse, as our priors from previous experiments would lead us to expect. This is important for our ability to directly interpret responses to the questions as subjective beliefs, as explained in the next section. The second reason was to measure the applicability of the Reduction of Compound Lotteries (ROCL) axiom for our sample. We explain later the fundamental importance of ROCL for the consequences of non-literacy, interpreted here as a non-degenerate subjective belief distribution in some domain.

The battery of lotteries we use was designed by Harrison, Martínez-Correa and Swarthout [2012] to allow for specific types of comparisons needed for testing ROCL over objective risk. Beginning with a given *simple* (S) lottery and *compound* (C) lottery, we create an *actuarially-equivalent* (AE) lottery from the C lottery, and then construct three pairs of lotteries: a S-C pair, a S-AE pair, and an AE-C pair. By repeating this process 15 times, we create a battery of lotteries consisting of 15 S-C pairs shown in Table A1, 15 S-AE pairs shown in Table A2, and 10 AE-C pairs[9] shown in Table A3. For transparency we used the universal "double or nothing" construction to generate compound lotteries.

Subjects were presented the 40 lottery pairs in random order, and selected the preferred lottery in each pair. Each lottery was displayed in a "pie chart" showing all probabilities and outcomes. All random draws were undertaken physically with dice. Detailed instructions and screen displays are provided in Appendix A.

[9] The lottery battery contains only 10 AE-C lottery pairs because some of the 15 S-C lottery pairs shared the same compound lottery.

2. Belief Elicitation

The decision maker in our experiment reports her subjective beliefs with a discrete version of a Quadratic Scoring Rule for continuous distributions, developed by Matheson and Winkler [1976]. Partition the domain into K intervals, and denote as r_k the report of the density in interval $k = 1, ..., K$. Assume for the moment that the decision maker is risk neutral, and that the full report consists of a series of reports for each interval, $\{ r_1, r_2, ..., r_k ,..., r_K \}$ such that $r_k \geq 0 \; \forall k$ and $\sum_{i = 1...K} (r_i) = 1$.

If k is the interval in which the true value lies, then the payoff score is from Matheson and Winkler [1976; p.1088, equation (6)]:

$$S = (2 \times r_k) - \sum_{i = 1...K} (r_i)^2$$

The reward in the score is a doubling of the report allocated to the true interval, and a penalty that depends on how these reports are distributed across the K intervals. The subject is rewarded for accuracy, but if that accuracy misses the true interval the punishment is severe. The punishment includes all possible reports, including the correct one.

Consider some examples, assuming $K = 4$. What if the subject has very tight subjective beliefs and puts all of the tokens in the correct interval? Then the score is

$$S = (2 \times 1) - (1^2 + 0^2 + 0^2 + 0^2) = 2 - 1 = 1,$$

and this is positive. But if the subject has a tight subjective belief that is wrong, the score is

$$S = (2 \times 0) - (1^2 + 0^2 + 0^2 + 0^2) = 0 - 1 = -1,$$

and the score is negative. So we see that this score would have to include some additional "endowment" to ensure that the earnings are positive.[10] Assuming that the subject has a very diffuse subjective belief and allocates 25% of the tokens to each interval, the score is less than 1:

$$S = (2 \times ¼) - (¼^2 + ¼^2 + ¼^2 + ¼^2) = ½ - ¼ = ¼ < 1.$$

[10] This is a point of practical behavioral significance, but is not important for the immediate theoretical point.

The tradeoff from the last case is that one can always ensure a score of ¼, but there is an incentive to provide less diffuse reports, and that incentive is the possibility of a score of 1.

To ensure complete generality, and avoid any decision maker facing losses, allow some endowment, α, and scaling of the score, β. We then get the generalized scoring rule

$$\alpha + \beta \, [\, (2 \times r_k) - \Sigma_{i=1 \ldots K} (r_i)^2 \,]$$

where we initially assumed α=0 and β=1. We can assume different values of α and β to transform the payoffs to any alternative range of levels we may want.

In our experiment K = 10, and we do not know whether the subject is risk neutral. Indeed, the weight of evidence from past experiments clearly suggests that subjects will be modestly risk averse over the prizes they face. It is well-known that risk aversion can significantly affect inferences from applications of the Quadratic Scoring Rule when eliciting subjective *probabilities* over *binary* events (Winker and Murphy [1970], Kadane and Winkler [1988]), and there are various methods for addressing these concerns.[11] Harrison, Martínez-Correa, Swarthout and Ulm [2012] characterize the implications of the general case of a risk averse agent when facing the QSR and reporting subjective *distributions* over *continuous* events, and find, remarkably, that these concerns do not apply with anything like the same force. For empirically plausible levels of risk aversion, one can reliably elicit the most important features of the latent subjective belief distribution without undertaking calibration for risk attitudes.

Specifically, they draw the following conclusions:

1. An individual reports having a positive probability for an event only if he has positive subjective probability for the event. So if the individual believes that unemployment is definitely below 12%, we would never see the individual reporting that it could be above 12%.

[11] For instance, see Köszegi and Rabin [2008], Holt and Smith [2009], Karni [2009] and Andersen, Fountain, Harrison and Rutström [2010].

Further, we can infer from Figure 1, for instance, that this subject truly attaches zero weight to the possibility of unemployment above 12%, no matter what his risk attitudes.

2. If an individual has the same subjective probability for two events, then the reported probabilities for the two events will also be the same if the individual is risk averse or risk neutral. So if the individual has a true, latent subjective probability of 0.1 that the unemployment rate is between 6% and 7.99%, and a true, latent subjective probability of 0.1 that it is between 10% and 11.99%, then the reported probabilities for these two intervals will be the same as well, as in Figure 1 (although typically not 0.1).

3. The converse is true for risk averse subjects, as well as for risk lovers. That is, if we observe two events receiving the same reported probability, we know that the true probabilities are also equal, although not necessarily the same as the reported probabilities.

4. If the individual has a *symmetric* subjective distribution, then the reported mean will be *exactly* the same as the true subjective mean, whether or not the subjective distribution is unimodal. Hence if we simply assume symmetry of the true distribution, a relatively weak assumption in some settings, we can elicit the mean belief directly from the average of the reported distribution.

5. The more risk averse an agent is, the more the reported distribution will resemble a uniform distribution defined on the support of their true distribution. In effect, risk aversion causes the individual to report a "flattened" version of their true distribution, but never to report beliefs to which they assign zero subjective probability.

6. It is possible to derive the effect of increased risk aversion on the difference between the reported distribution and true distribution. Harrison, Martínez-Correa, Swarthout and Ulm [2012] show numerically that *a priori* plausible levels of risk aversion in laboratory settings implies no significant deviation between reported and true subjective beliefs in this setting.

Provided that our subjects exhibit the modest levels of risk aversion that are typically found in lab settings with similar stakes, these results provide the basis for using the reported distributions as if they are the true, subjective belief distributions.

3. Results on the Measurement of Literacy

A. Description of Results

In December 2012 we recruited 120 subjects from Durham University. The majority had major fields of study in Finance or some other Business area, and were completing a Master of Science degree. The average age was 24.4, 67% were women, and 85% were single and had never been married. Just over 73% were non-EU citizens, and 14% were current smokers.

Figures 3 and 4 provide a quick helicopter tour of the aggregate beliefs we elicited. More formal statistical tests will be provided below. We observe very precise beliefs for the interest compounding question, which was relatively easy for our sample. Far less precision is observed for the other domains. Aggregate beliefs for the economic literacy questions tended to be unimodal, with most subjects having some sense of where the correct answer was, but with varying precision as we will see. But when we come to the statistical literacy questions we see some interesting bimodal aggregate beliefs. With a few exceptions, to anticipate, this bimodality comes from different individuals: it is not the case that the same individual, in general, held such disparate views. The suggestion, then, is that some subjects "got it," and some subjects "had no clue."[12]

Figures 5 and 6 begin the evaluation of individual responses for the two financial literacy questions about interest compounding and inflation and the value of money. In each case we report

---

[12] We are, of course, not the first to point out that incentivized subjects have problems applying Bayes Rule. See Grether [1992] and Holt and Smith [2009] for careful results, and references to the larger literature. El-Gamal and Grether [1995] go further and elegantly propose a finite mixture model to classify subjects as exhibiting different models of updating.

the correct answer, a "literacy index" and "concordance index," the responses of three individuals selected to illustrate some differences in individual behavior, and the pooled distribution. We discuss the concordance index in the next section.

We construct a simple index of literacy, $L \in [0, 1]$, given by the fraction of 100 tokens that the individual allocates to the interval containing the true answer. This index does not need to be estimated: it is a direct transformation of the observed data. Thus we see a value of $L = 1$ for subject #1 in Figure 5, the interest compounding domain, since this subject allocated all 100 tokens to the interval containing the correct answer. Many subjects did exactly the same thing in this case, but subject #10 and subject #11 show how a few hedged their bets, quite literally. For subject #10, 60 of the 100 tokens were in the interval containing the correct answer, so $L = 0.60$ for this subject.

In Figure 6 we see that subject #1 has a literacy index of zero, since she allocated all 100 tokens to the interval just to the left of the correct answer for the inflation question. In this domain, subject #5 has a literacy index of 0.26 since 26 out of 100 tokens were allocated to the correct interval. By being less dogmatic, subject #5 exhibited greater literacy than subject #1. Of course, one might want to argue that subject #1 was very close to the correct answer, but countering that claim is the subject's choice, implicitly saying that she was certain of her answer. If indeed she has some imprecision, that should have led her to report a non-degenerate distribution.

Figures 7 and 8 consider the other economic literacy questions, about expected remaining lifetime for men and women. Figure 7 shows detailed responses for 11 individuals to the question about men, since we observe considerable heterogeneity in this domain compared to the financial literacy questions. The imprecision for subject #3, #5 and #7 is substantial, and leads one to speculate if it shows up in their savings behavior or retirement planning. Figure 8 shows the differences in the aggregate distribution between the questions about men and women, to gauge if aggregate literacy detects the longer expected lifetime of women compared to men. Indeed, we see that this increment is detected.

Turning to the statistical risk questions in Figures 9 through 12, we uncover some dramatic differences in individual literacy.

Consider the responses in the breast cancer risk domain, shown in Figures 9 and 10. Figure 9 displays 11 individual belief distributions, and Figure 10 then shows the pooled differences between using the probability format for conveying information and the frequency format. There are many subjects who exhibit considerable statistical literacy with this question: subjects #1, #3, #4 and #6 are fully literate, and subjects #2, #10 and #11 exhibit just minor uncertainty. But subjects #5, #7, #8 and #9 illustrate a pattern that signifies an inability to process the information provided (or, to be sure, an inability to understand how it translates into responses).

Is this striking heterogeneity of statistical literacy improved by using a frequency format? Figure 10 shows that it is, on balance. However, there remain a large number of individuals who exhibit some beliefs that are statistically incorrect.

Turning finally to the credit risk domain, we see a similar pattern of heterogeneous literacy. Several of the high-literacy subjects in the breast cancer domain also exhibited high levels of literacy in this domain (e.g., #1, #3, #4, #6 and #10), and several of the low-literacy subjects did similarly in this domain (e.g., #5, #7, #8 and #9). So the pattern that emerges from Figures 9 and 11 is of individuals who are sharply divided between those who are highly-literate and those who have no clue about the correct statistical inference. This is different from the uncertainty that each individual generally had with respect to the question about the remaining lifetimes of men and women. In contrast to the finding for breast cancer risk, Figure 12 also suggests that there was no improvement in statistical literacy in the credit risk domain as we moved from using probability information to frequency information in the statement of the question.

Finally, Figures 13 and 14 collate information on the distribution of the literacy indices *L* across the domains considered. The vertical, dashed line is the average of each distribution, for reference. The distribution for the interest compounding question, in the top left panel of Figure 13, is what one would normatively like to see: almost universal high-literacy. However, one can visually infer that this is the exception across these economic and statistical domains. The distributions for statistical literacy might seem inconsistent with the displays of pooled beliefs, which show the largest mode at the correct response. If this is the case, how can the largest mode for the literacy index be at zero? The answer is that the subjects who understood how to draw statistical inferences tended to do so with certainty, allocating all 100 tokens to the correct interval, and the subjects who did not understand how to respond allocated tokens all over the possible range. There is one way to be completely correct about these questions, and 4,263,421,511,270 ways to be wrong.[13]

B. Statistical Analysis of Results

Before the beliefs data are analyzed, we must assess whether the risk attitudes of individuals are within the limit that leads to no significant incentive for subjects to distort their true beliefs. Applying maximum likelihood methods described by Harrison and Rutström [2008] to the lottery choice data pooled over all individuals, we estimate a relative risk aversion of 0.74 with a 95% confidence interval between 0.69 and 0.78. These estimates are well within the tolerance that Harrison, Martínez-Correa, Swarthout and Ulm [2012] suggest (only values as high as 2 or 3 start to cause problems of interpretation). We also estimate relative risk aversion values for each individual, and the distribution of these estimates is also well within the range that is required.[14]

---

[13] If there are t tokens and b bins, then there are (t+b-1)!/t!(b-1)! possible allocations in each of our elicitation tasks. Only one of these is completely correct. If someone has a literacy index $L = 0$ then there are still (100+9-1)!/100!(9-1)! = 352,025,629,371 ways to respond.

[14] Figure B1 in Appendix B displays this distribution. Numerically reliable estimates are obtained for 114 of 120 subjects.

A natural statistical model for directly evaluating the beliefs data is interval regression. In this specification the dependent variable refers to the intervals given by our elicitation "bins."[15] In all cases we control for sex, age, marital status, race, whether a Finance major, whether a non-EU citizen, whether a current smoker, and the scores on the CRT and Berlin Numeracy Test. The age and test scores are all normalized to have mean zero and unit standard deviation. Detailed estimates are provided in Appendix B.

The interval regression models for each belief question[16] show the following statistically significant demographic effects:

- Older individuals are slightly more literate in the interest compounding domain, although there is very little variation in the dependent variable here (Figure 5).
- Women have slightly lower literacy in the inflation domain.
- Whites are much more literate on the expected remaining life years for men (and also for women, although the effect for women is not statistically significant).
- Older individuals are substantially more literate on the expected remaining life years for women.
- Finance majors and those scoring well on the Berlin Numeracy Test were much more literate on the breast cancer risk question when presented in probability format, but whites and current smokers are clearly identifiable as those that did not "get it."
- The effects of demographics for the breast cancer risk question when presented in frequency

[15] Interval regression allows one to identify clopen intervals with -∞ as a lower bound or +∞ as an upper bound, but that is not generally appropriate for our elicitation tasks. We evaluate the effects of using such clopen bounds for the interest compounding and inflation questions, and it makes no difference to our conclusions.

[16] For the breast cancer risk and credit risk questions we only consider estimates for the sample that got each indicated format as the first question (Tables B7 and B8). There appears to be a rich demographic interaction with these order effects of the format, as suggested by the contrasting estimates for the same characteristic in Tables B5 and B7, and in Tables B6 and B8.

format were in the same direction as with the probability format, but none were statistically significantly different from zero.

- Single, older individuals and those who scored well on the Berlin Numeracy Test were more literate on the credit risk question in probability format.
- Only older individuals were more literate on the credit risk question in the frequency format.

We evaluate the systematic effect of demographics more formally below by pooling across measures of literacy L in different domains.

The interval regression model for the two questions on breast cancer risk,[17] confirms the inferences drawn from Figure 10. Using the frequency format to present the basic information needed to apply Bayes Rule improves literacy by moving average responses closer to the true value by 8.8 percentage points, and the estimate of this effect has a p-value of 0.001. Being a Finance major also improves literacy in this domain, by 8.8 percentage points (p-value of 0.078). And the Berlin Numeracy Test score is also correlated with a significant improvement in statistical literacy of 7.2 percentage points (p-value of 0.012). The CRT score has no statistically significant effect on responses. Reductions in literacy are associated with being White (14.5 percentage points, and a p-value of 0.010) and a current smoker (16.6 percentage points, and a p-value of 0.018).[18]

However, when applied to the credit risk responses, we see no effect from using the frequency format.[19] Average beliefs are actually lowered by the frequency format, *away* from the true posterior probability, by 1.1 percentage points, but this effect is not statistically significantly different from zero (p-value of 0.66). The Berlin Numeracy Test score again is associated with an improvement in literacy in this statistical task: responses increase with 4.6 percentage points (p-value of 0.03). And being single

---

[17] Table B13.

[18] The same qualitative results are obtained if we restrict the sample to those who were presented the frequency format as the second question (Tables B14).

[19] Table B15.

is associated with a better literacy in this domain (by 15.1 percentage points, with a *p*-value of 0.025). Finally, women exhibit significantly poorer literacy in this domain (by 9.9 percentage points, with a *p*-value of 0.02).[20]

To assess the effects of demographics on literacy we estimate an ordered logit model of the literacy index L, but collapse the index to three values for ease of interpretation. One value corresponds to $0 \leq L \leq ⅓$ and might be called "illiterate," another value corresponds to $⅓ \leq L \leq ⅔$ and might be called "semi-literate," and a final value corresponds to $L > ⅔$ and might be called "literate." Table 1 reports the marginal effects[21] of the listed covariates on the probability of being illiterate, semi-literate and literate, as defined here.

These effects are normalized to the high literacy found in the interest compounding question, so it is not surprising to see most of the dummies for individual questions showing an increased effect on the probability of illiteracy by comparison. For instance, compared to the interest compounding question, the effect of asking the inflation question was to increase the probability of illiteracy ($L \leq ⅓$) by 0.39. We observe a large, positive, statistically significant effect of +0.11 on the probability of being literate ($L > ⅔$) from asking the breast risk question with the frequency format, but a small, negative, statistically significant effect of -0.05 from asking the credit risk question with the frequency format.

Focusing only on statistically significant effects, in this sample women are more likely to be illiterate and less likely to be literate than men. Older subjects are more likely to be semi-literate and literate. Marriage is associated with a striking increase in illiteracy. Being a Finance major is associated with heightened literacy. Being a non-EU citizen is associated with a much lower level of literacy

[20] The same qualitative results are obtained if we restrict the sample to those who were presented the frequency format as the second question (Tables B16).

[21] These are the *average* marginal effects, evaluated by considering all of the actual values of the non-target variables and averaging the marginal effects across those values. The alternative is to consider the average values of the non-target variables and reporting the marginal effect at that single set of values.

compared to EU citizens. Finally, the Berlin Numeracy Test is associated with a significantly higher probability of being literate. These demographic effects are a mix of the expected (e.g., the Berlin Numeracy Test) and the unexpected (e.g., being single). The significance of these demographic effects also points to the heterogeneity of literacy across these domains.

4. Results on the Consistency of Knowledge

Although literacy is a capacity that is naturally measured for the individual, it obviously impacts the extent to which knowledge about something is shared. We consider the idea of social literacy later, but if someone has a poor level of literacy in some domain, the natural question is whether that is consistent with the knowledge that others have. The immediate consequences for behavior when there are heterogeneous beliefs are by now well-studied, such as in models of asset pricing in finance (e.g., Shefrin [2008]), game theoretical interaction, and rational expectations.

These ideas are also familiar from linguistics. The process of learning a language involves the disambiguation of utterances (Allen [1995]). And many linguists discuss language use as intentionally constrained by norms of communicating understanding, which is to say greater literacy (Grice [1989], Clark [1992]). Hence one naturally seeks some measure of shared literacy. Is the uncertainty over some fact in a given domain shared, or is it a domain in which one can clearly identify "experts" and "novices?" We propose a simple measure that can allow us to address that question.

Any measuring instrument can be compared against another measuring instrument. Examples include weight scales, political opinion polls, or medical judgements about diagnoses. In our case we consider the subjective beliefs about some fact, and seek to measure their consistency. In the biostatistics literature a popular concordance index ϱ_c has been developed by Lin [1989][2000]. It combines the familiar notion of correlation from a Pearson inter-class correlation coefficient with allowance for bias, and is virtually identical to measures of intra-class correlation (Nickerson [1997]). It is bounded between ±1, with the usual interpretation that $\varrho_c = 1$ indicates perfect concordance, and smaller values indicate poorer concordance.

In Figure 5, for instance, we evaluate the concordance index for each subject with respect to the pooled belief distribution on the interest compounding question, and then also report the average value of the index over all 120 subjects. Even though subject #1 in that setting had a literacy index value of 1, since she gave the correct responses, her concordance with the group was slightly less than 1 (0.972) because some people in the group did not have perfect literacy in this domain (e.g., subjects #10 and #11, shown in Figure 5). Taking a less extreme case, such as the inflation and value of money question in Figure 6, we see much lower levels of concordance. Subject #5, even though less precise than subject #1 and subject #3, was more consistent with the beliefs that everyone else had.

Moving to the distribution of concordance indices in each domain, Figures 15 and 16 show the heterogeneity of beliefs we elicited. The interest compounding question, in Figure 15, is again an outlier, showing considerable literacy in this sample (from Figure 13) and hence considerable consistency.

5. Limitations and Extensions

A. Capability

The modern policy literature on literacy stresses the concept of "capability," which is the extent to which individuals use their knowledge, as distinct from being able to answer abstract questions successfully. The concept of capability seeks to characterize if someone is able to function in a certain domain. This raises many subtle, interesting issues.

It is not obvious that someone must know the right answer in order to be able to function in some domain robustly. There are some task domains where the payoffs are very "flat," in the sense that large errors in the specific choice lead to virtually the same expected payoff as more refined choices. These are well known in experimental economics (Harrison [1989][1992][1994]).

It is also not obvious that someone must infer the right answer by applying grammatical or logical algorithms in order to make good choices: heuristics might do very well in many domains, whether or not there is a flat payoff at work in the region of choice. This is a major theme of the work of Gigerenzer [1996], and runs counter to the presumption that many draw from Kahneman and Tversky [1996] that heuristics generally lead to sub-optimal choices.

The concept of capability also raises the issue of domain-specific knowledge, which goes beyond the "reasoning from first principles and the information in front of you" approach that characterizes most analyses of literacy. Someone might be a wizard at applying Bayes Rule, but simply have an incorrect prior belief about some base rate. Such a person would typically be deemed statistically literate but not capable.

The notion of capability can be directly evaluated by adding tasks that apply some of the knowledge about which someone is supposed to exhibit literacy. A good illustration in the context of hypothetical survey evidence on inflation expectations is provided by Armentier, de Bruin, Topa, van der Klaauw and Zafar [2011], who pose an investment task to subjects that depends, in part, on their expectations of inflation. Do people act on their expectations and literacy? Our incentivized elicitation procedures build in a check for that, of course, since the subjective beliefs are revealed to us through structured bets that involve real stakes. But the question is whether or not the knowledge demonstrated in one literacy task is correctly applied in related contexts. For instance, does the degree of literacy that our subjects exhibit in Q2 on the distinction between nominal and real values affect their propensity to suffer from a "money illusion" in an experiment of the kind reported by Fehr and Tyran [2001]?

More generally, the concept of financial literacy is not just the same as financial knowledge, although the latter is obviously an important component of the former. Remund [2010] reviews the many definitions of financial literacy found in the academic and policy literature, and correctly notes that operational measures need not entail knowledge. For instance, someone might hire an expert to complete their taxes or design their retirement portfolio, and not know anything about tax rules or retirement planning. Of course, one might say that this reflects "embodied" or "social" knowledge, rather than individual knowledge, but that in itself is an important distinction we return to below.

Huston [2010] conducts a meta-analysis of research into financial literacy, and noted that 47% of the 71 studies evaluated used the terms financial literacy and financial knowledge synonymously. Indeed, if one just looks at the 62% of studies that referred to both, over 75% used them interchangeably. She correctly concluded that "If these two constructs are conceptually different, then using the terms interchangeably indicates a potential problem" (p. 303).

B. The Consequences of Non-Literacy

If someone exhibits a low degree of literacy in some domain, does it follow that he will act on his beliefs in the same manner as someone who exhibits a higher degree of literacy? In other words, if one person knows that he does not have precise knowledge in a certain domain, does he act on that knowledge of his relative literacy? This question goes to the heart of distinctions between risk, uncertainty and ambiguity.

If somebody has a well-defined subjective belief distribution about some matter, Subjective Expected Utility (SEU) assumes that they act as if they reduce that distribution down to it's weighted mean, and then make a decision as if that weighted mean probability was known with certainty. To illustrate, assume a three-point discrete, non-degenerate, subjective distribution over a binary event in which the individual holds subjective probability $\pi = 0.6$ with "prior" probability 0.1, $\pi = 0.7$ with "prior" probability 0.6, and $\pi = 0.8$ with "prior" probability 0.3, for a weighted average $\pi = 0.72$. Now consider a lottery in which one gets \$X if the event occurs, and \$x otherwise. Then the SEU is

$$0.1\times0.6\times U(X)+0.1\times0.4\times U(x)+0.6\times0.7\times U(X)+0.6\times0.3\times U(x)+0.3\times0.8\times U(X)+0.3\times0.2\times U(x),$$

which collapses to

$$(0.1\times0.6 + 0.6\times0.7 + 0.3\times0.8) \times U(X) + (0.1\times0.4 + 0.6\times0.3 + 0.3\times0.2) \times U(x),$$

and hence to

$$0.72 \times U(X) + 0.28 \times U(x)$$

under the ROCL axiom. So the original non-degenerate distribution can be reduce down to a degenerate subjective probability of 0.72 under ROCL: an impressive identifying restriction.

Now return to our question about the differential behavior of someone who is less literate in some domain, and knows it. Being "less literate" does *not* mean that one has zero mean errors on either side of the true value. Hence it is far from obvious that there is any normative appeal to applying ROCL in this setting. Indeed, common sense indicates that the decision-maker, armed with knowledge of his lack of literacy, should somehow take that lack of precision in beliefs into account. In other words, the decision maker should, in some manner, violate ROCL.

Figure 17 provides some insight into the practical significance of this point. By replacing the subjective belief distributions generated by each of our subjects on each of the 8 questions with their weighted mean and then recalculating the literacy index L with this SEU-consistent probability, we see the effect of being SEU-consistent on literacy. Of course, one corollary is that this SEU-consistent measure of literacy is binary: one either has a weighted mean in the true interval or not. Averaging over all 120 individuals, this SEU-consistent literacy index is generally lower than the raw literacy index L. It is not dramatically lower for the statistical questions, or the interest compounding question that virtually every subject answered correctly. But for some domains, such as understanding the effects of inflation and knowing expected remaining lifetime, the implications of applying SEU are a dramatic reduction in apparent literacy. For a few individuals, the application of ROCL increases the literacy index L for some domains; this occurs when their weighted mean was correct, but they had assigned some subjective belief to other values.

How we relax ROCL is a matter for important, foundational research. Although it has taken half a century for the implications of Ellsberg [1961] to be formalized in tractable ways, we are much closer to doing so. One popular approach is the "smooth ambiguity model" of Klibanoff, Marinacci and Mukerji [2005], although there are many other competing specifications in the literature.

To provide a concrete example, we illustrate the smooth ambiguity model with some simple numbers. Let $CE(\pi=0.6)$ be the certainty equivalent of the lottery $0.6\times U(X)+0.4\times U(x)$, $CE(\pi=0.7)$ be the certainty equivalent of the lottery $0.7\times U(X)+0.3\times U(x)$, and $CE(\pi=0.8)$ be the certainty equivalent of the lottery $0.8\times U(X)+0.2\times U(x)$. Using the priors from our previous example, the evaluation of the lottery is

$$0.1\times\varphi(U(CE(\pi=0.6))) + 0.6\times\varphi(U(CE(\pi=0.7))) + 0.3\times\varphi(U(CE(\pi=0.8))),$$

where φ is a function defined over the domain of U(.). Akin to the properties of $U(\cdot)$ defining risk attitudes under SEU, the properties of $\varphi(\cdot)$ define attitudes towards the uncertainty over the particular subjective probability value.[22] If φ is concave, then the decision-maker is uncertainty averse; if φ is convex, then the decision-maker is uncertainty loving; and if φ is linear, then the decision-maker is uncertainty neutral. The familiar SEU specification emerges if φ is linear, since then ROCL applies after some irrelevant normalization. The overall evaluation of the lottery depends on risk attitudes *and* uncertainty attitudes, and there is no reason for the decision-maker to be averse to both at the same time.

We can now restate our initial question yet again: is someone who exhibits less literacy in some domain also likely to exhibit violations of ROCL in that domain? Although we generally want to avoid

---

[22] In the original specifications φ is said to characterize attitudes towards ambiguity, but the earlier definition of risk, uncertainty and ambiguity makes it apparent why one would not want to casually confound the two. One would only be dealing with ambiguity in the absence of well-defined prior probabilities over the three subjective probability values 0.6, 0.7 and 0.8.

the expression "illiteracy" in lieu of "degree of literacy," this restatement is in effect equating "illiteracy aversion" with "uncertainty aversion" or "ambiguity aversion."

We are a long way from being able to rigorously test if someone behaves consistently with ROCL defined over subjective belief distributions.[23] But it is relatively straightforward to gather evidence on a closely related hypothesis concerning the violations of ROCL defined over *objective* risk, which was the secondary purpose of our lottery choice tasks. Figure 18 shows a distribution of *p*-values of hypothesis tests of ROCL for each individual. This test is constructed by assuming that the individual has a different risk attitude for simple lotteries and compound lotteries, estimating those two risk attitudes using maximum likelihood methods, and then testing for the equality of those two risk attitudes.[24] Those in the mode to the left of Figure 18 are classified as violating ROCL, and those to the right of that mode are classified as behaving consistently with ROCL. We stress, again, that these *p*-values test ROCL over objective probabilities, not subjective beliefs.

Figure 19 then collects the information in Figure 17 and Figure 18 for each individual. The left panel of Figure 19 shows the relationship between the literacy index *L* for each individual, averaged over all 8 domains, and their *p*-value on the test of ROCL. Normatively, one would have hoped to see

[23] Indeed, the identification problems run deeper. It is tempting to argue that the elicited subjective distributions reveal what Ghirardato, Maccheroni and Marinacci [2004] refer to as "revealed ambiguity," which is the set of priors that a person has about some events. This allows us, following them, to separate the individual's perceived ambiguity from that individual's reaction to it: for instance, if ROCL is applied, as we just argued, then the individual is ambiguity-neutral. But then we have to recognize that any given stimulus or domain might be ambiguous to one person and not ambiguous to another person. This is anticipated clearly by Ghirardato, Maccheroni and Marinacci [2004; p. 136] who note that "... nothing precludes two DMs from perceiving different ambiguity in the same decision problem." The implication of this insight is that experimental economists cannot just confront subjects with a task, such as the Ellsberg two-urn task, and assume that every subject perceives it as ambiguous. The fact that such tasks are conventionally labeled as "ambiguous" in the literature does not make them so in the minds of subjects. One non-ambiguous interpretation in this case is that the experimenter picked the fraction of red and black balls in the "ambiguous" urn using a uniform distribution over all outcomes. This interpretation removes all ambiguity, and reduces the problem to a compound lottery over objective risk (which could still be evaluated with or without the help of ROCL).

[24] This is the "source-dependent" model of Harrison, Martínez-Correa and Swarthout [2012]. Estimation was undertaken at the level of the individual. Out of 120 subjects, 6 did not numerically converge and are dropped from this analysis.

a tendency for individuals to *violate* ROCL, since applying ROCL tends, from Figure 17, to increase overall literacy. If anything we see a slight tendency for a positive relationship: a descriptive Tobit regression of the literacy index L on the p-values testing ROCL has a coefficient of +0.12 with a p-value itself of only 0.072. So there is a very slight tendency for those with higher literacy indices to behave as if they are ROCL-consistent over objective lotteries. The right panel of Figure 19 examines this issue in a more direct manner, by looking at the relationship between the *reduction* in literacy for an individual from applying ROCL and their p-value on the test of ROCL. Here we would have normatively hoped for those with big reductions in literacy from applying ROCL, big positive values in this scatter plot, to be the ones for whom ROCL was rejected. This would imply a negative relationship in this panel of Figure 19, and clearly there is none.[25] Again, our measure of ROCL is over objective lotteries, and one can imagine a pun-loving subject saying, "I don't have a problem with compound objective lotteries, or even with easy compounding questions, but I will certainly be careful applying ROCL in domains like statistical inference or harder economic choices such as my remaining lifetime."

Taken together, Figures 17, 18 and 19 show the significance of understanding the relationship between literacy and ROCL, if one is to understand the consequences of non-literacy for choice. In the spirit of the "second best," it is not at all obvious that obeying ROCL is normatively attractive if one has subjective beliefs that are not, on average, correct.

C. Social Literacy

When measuring the literacy of a household, how does one account for the heterogeneity of levels of literacy within the household? The concept of effective literacy, developed by Basu and

[25] A descriptive regression of the reduction in the literacy index L on the p-values testing ROCL has a coefficient of 0.042 with a p-value of 0.30.

Foster [1998], considers this important dimension of what might also be called "social literacy." Obviously the measurement issues goes beyond the household, and includes any social network used by individuals for making decisions.

One approach focuses on potential literacy, and defines literacy in terms of the most literate person in the household. This adjustment to naïve measures of literacy at the individual level is easy to make, and provides a valuable upper bound. One could, of course, similarly define a lower bound if the household power relationships lead to the least literate person imposing his or her will on the household decision in some setting. In general, social literacy is a public good, and might be subject to varying "composition functions" in the sense of Harrison and Hirshliefer [1989]. The upper and lower bounds just noted reflect "best shot" and "weakest link" composition functions, but of course there could be a wide range of (less extreme) functions, and they could vary from domain to domain. Using our tools for characterizing the production function for social literacy in these settings is an important extension.

D. Hypothetical Surveys and the Measurement of Literacy

Most surveys of literacy are hypothetical, in the sense that they do not pose incentivized questions as we do. An obvious question is whether this leads to any bias in responses, with the presumption to economists that having incentives will provide more reliable responses.[26] In fact, there are two components of these hypothetical questions: one is the lack of any financial or economic

[26] Delavande, Giné and McKenzie [2011; p. 156] make the case for not bothering about incentives. Referring to studies in developing countries that have all been hypothetical, they argue that "even without payment, the answers received from such questions appear reasonable, and as such, there seems to have been a *de facto* decision that payments are not needed." We do not know what "reasonable" might possibly mean when it comes to subjective beliefs. In some settings, such as stated beliefs about longevity (e.g., Perozek [2008]), the metric for reasonableness appears to be whether the beliefs are actuarially correct on average. Although that is certainly of great policy interest, it is hard to know why it would be a metric for evaluating the validity of responses as reflecting the true subjective beliefs of individuals.

consequence to giving one answer rather than another, and the other is whether or not incentives actually encourage truthfulness. It is easy to come up with scoring rules or prediction markets, for instance, that do not elicit responses that can be meaningfully interpreted even if there are financial rewards involved (e.g., Manski [2006] and Fountain and Harrison [2011]).

We expect there to be some correlation between well-constructed, popular, hypothetical surveys of literacy and our measures, in the sense that the former are likely to be statistically informative about the latter. Indeed, we envisage a complementarity between the two. Large samples can be collected using hypothetical surveys, and then calibrated using results from incentivized responses to the same questions from a smaller sub-sample drawn from the same population (e.g., Blackburn, Harrison and Rutström [1994]).

E. Measuring Literacy Without Assuming Literacy

An important and obvious question is how to design measures of literacy that do not presume certain kinds of literacy. This question becomes critical when one recognizes that informal knowledge can of course be held without the ability to formally communicate it by reading, particularly if the text involves jargon of some kind, such as probabilistic statements. An excellent example of the type of modifications that can be made to elicit responses is provided by Delevande and Kohler [2012; Online Supplement], who elicited hypothetical subjective belief distributions from individuals in Malawi. Although they did not use any scoring rule at all, their language explained the mechanism in simple terms:

> I will ask you several questions about the chance or likelihood that certain events are going to happen. There are 10 beans in the cup. I would like you to choose some beans out of these 10 beans and put them in the plate to express what you think the likelihood or chance is of a specific event happening. One bean represents one chance out of 10. If you do not put any beans in the plate, it means you are sure that the event will NOT happen. As you add beans, it means that you think the likelihood that the event happens increases. For example, if you put one or two beans, it means you think

> the event is not likely to happen but it is still possible. If you pick five beans, it means that it is just as likely it happens as it does not happen (fifty-fifty). If you pick six beans, it means the event is slightly more likely to happen than not to happen. If you put ten beans in the plate, it means you are sure the event will happen. There is no right or wrong answer, I just want to know what you think. Let me give you an example. Imagine that we are playing Bawo [a local board game with objective probabilities]. Say, when asked about the chance that you will win, you put seven beans in the plate. This means that you believe you would win seven out of ten games on average if we play for a long time.

Our approach, following Andersen, Fountain, Harrison and Rutström [2010], is to use visual, computerized interfaces that embody formal scoring rules "underneath the hood." These interfaces only present individuals with the implied lotteries from using those rules, and do not need to confront subjects with complicated payoff tables of numbers. We recognize the current limitations of these approaches for many populations of interest in the measurement of literacy.

6. Conclusions

Literacy is a concept that is widely discussed, and clearly at the core of understanding economic behavior in modern societies. We propose a characterization of literacy using the familiar notion of a subjective belief distribution over questions in a certain domain. We demonstrate how these belief distributions can be elicited in an operational, incentive-compatible manner. We show that there is considerable heterogeneity in literacy levels over economic, financial and statistical domains, and across observable demographics. We also demonstrate that uncertainty aversion might be normatively attractive to those with imperfect literacy.

The immediate extensions of our approach are to consider broader samples and other domains of literacy, as well as the effects of controlled interventions on the distribution of literacy. Evaluation of the consequences of imperfect illiteracy can be undertaken by studying the choices made, or avoided, in related tasks that rely on literacy in that domain. Do semi-literate individuals avoid welfare-improving choice domains for fear that they might make serious mistakes? Relatively straightforward extensions to consider social literacy are also important and natural using our characterization. More challenging is the evaluation of the extent to which individuals apply ROCL over subjective belief distributions, and hence whether uncertainty aversion is normatively appealing.

Figure 1: Belief Elicitation Interface

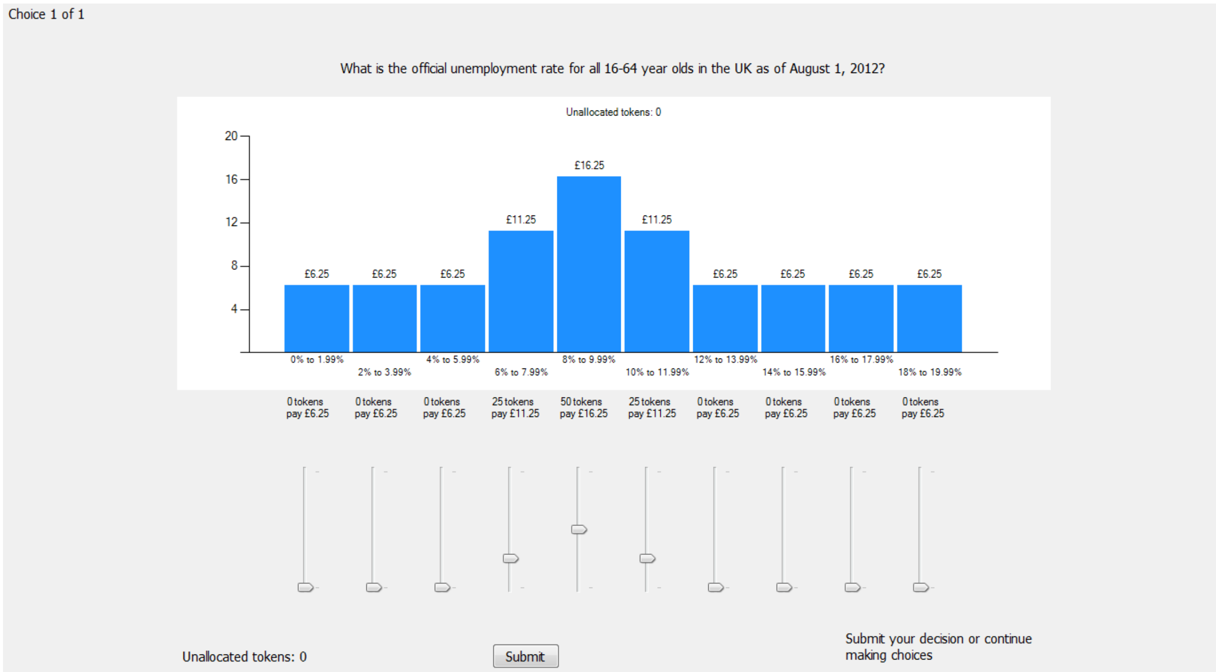


Figure 2: Possible Belief Elicitation Response

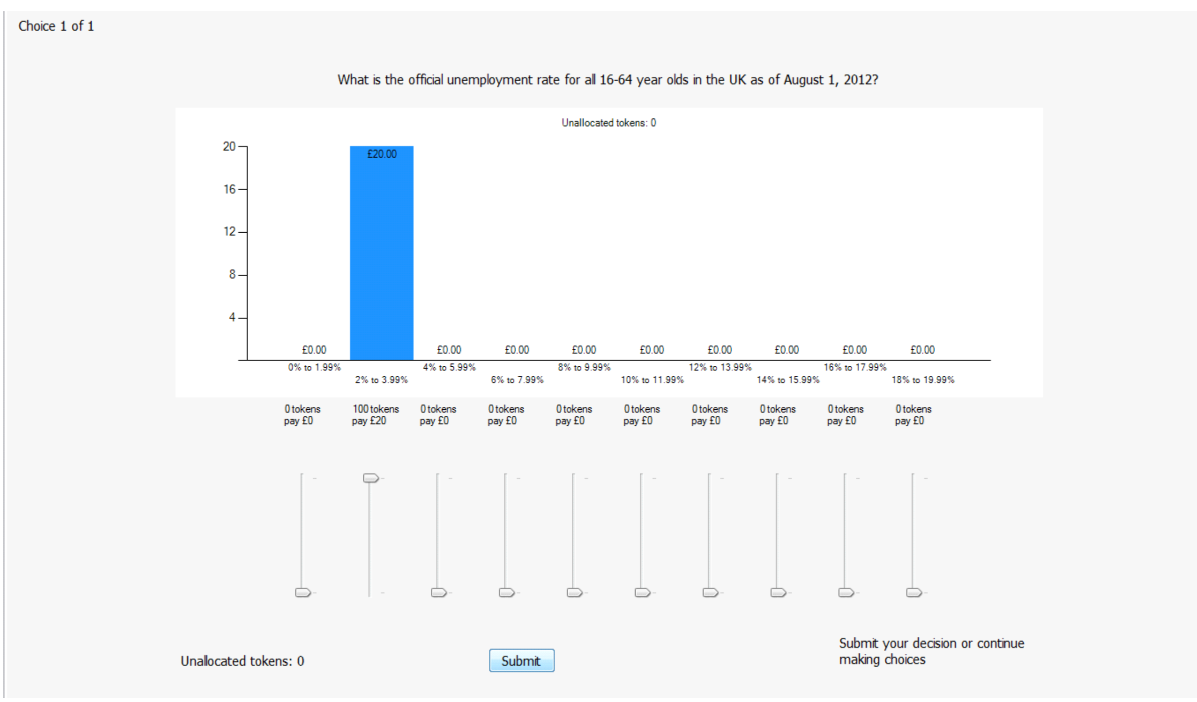


Figure 3: Pooled Subjective Beliefs for Economic Literacy Questions

Labels in left (right) panel show lower bound (midpoint) of elicited interval

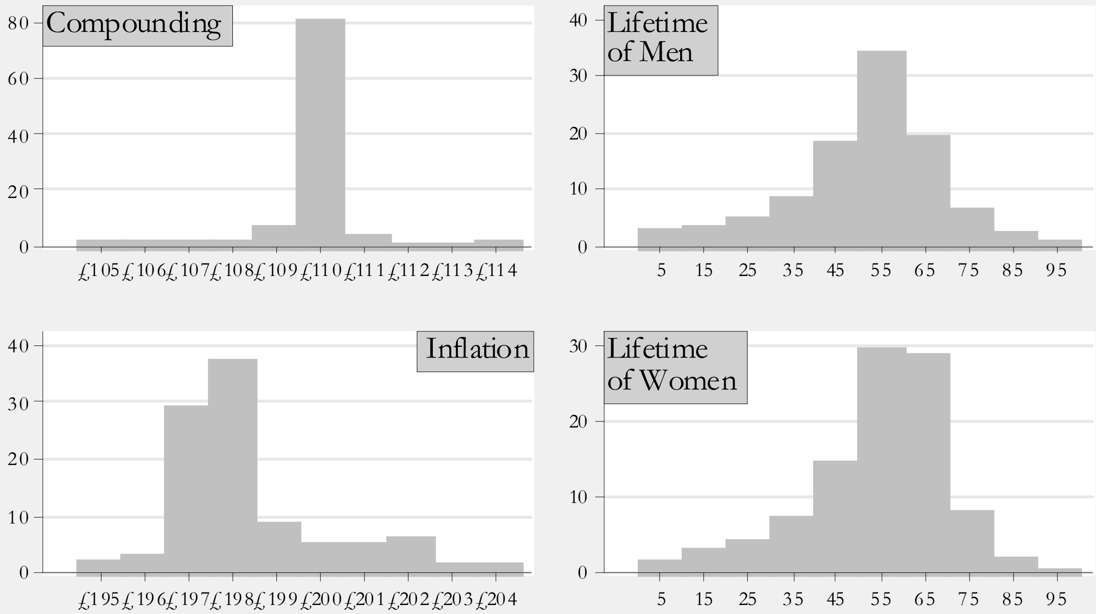


Figure 4: Pooled Subjective Beliefs for Statistical Literacy Questions

Labels show midpoint of elicited interval

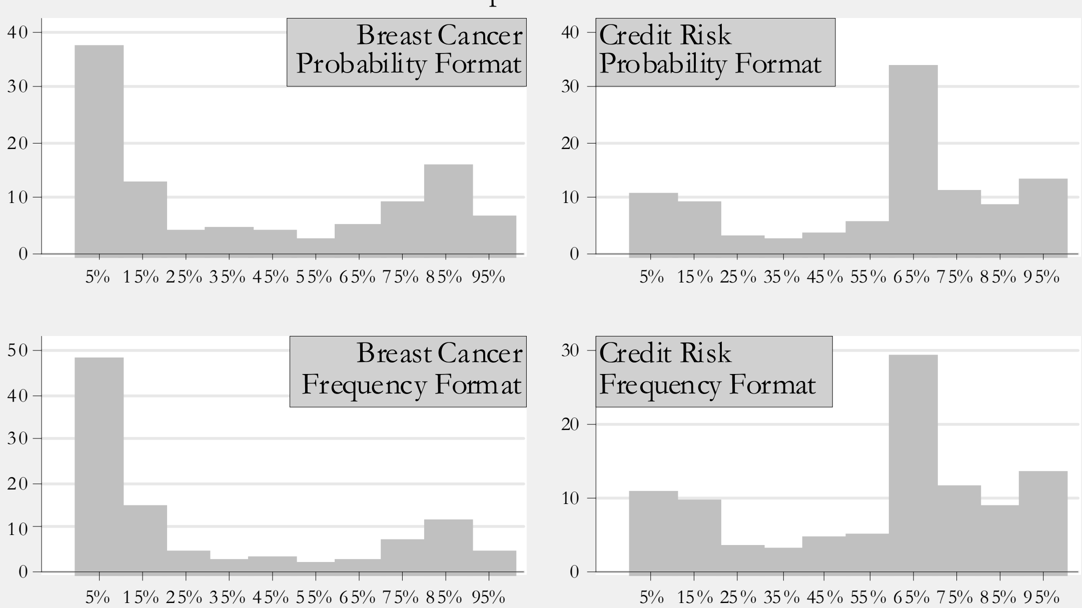


Figure 5: Subjective Beliefs of Three Subjects to Interest Rate Compounding Question

Correct Answer is £110.40

Average Literacy index L = 0.81 and average Concordance index ρ_c = 0.879

Labels show lower bound: e.g., £105 refers to the interval £105 to £105.99

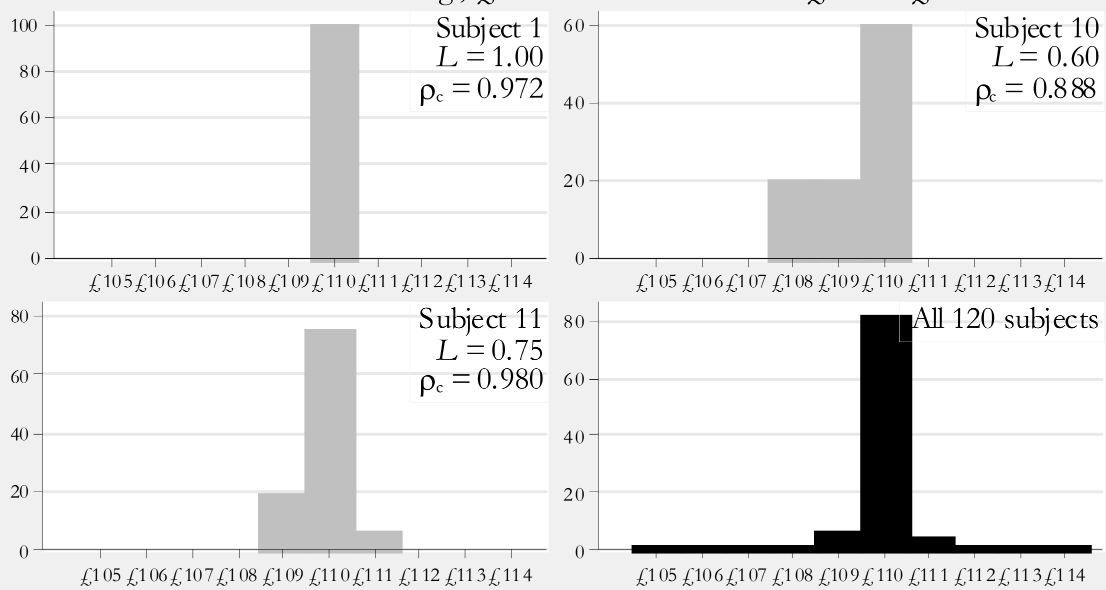


Figure 6: Subjective Beliefs of Three Subjects to Inflation and the Value of Money Question

Correct Answer is £198

Average Literacy index L = 0.38 and average Concordance index ρ_c = 0.434

Labels show lower bound: e.g., £195 refers to the interval £195 to £195.99

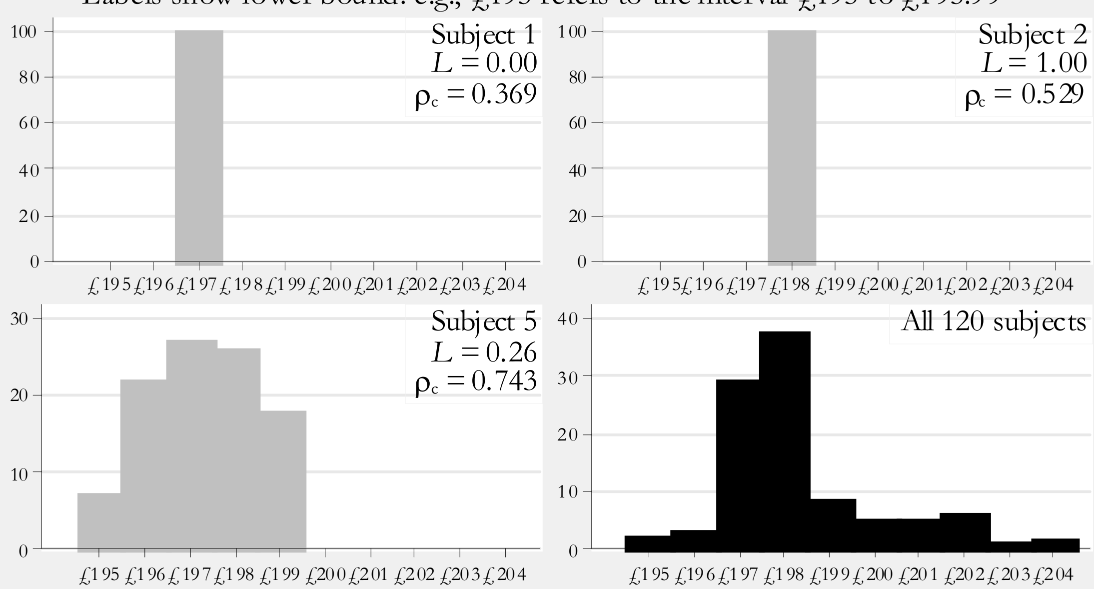


Figure 7: Beliefs on the Remaining Lifetime of Men

Correct Answer is 59.1

Average Literacy index $L = 0.34$ and average Concordance index $\rho_c = 0.523$

Labels show midpoint: e.g., 5 refers to the interval 0 to 9

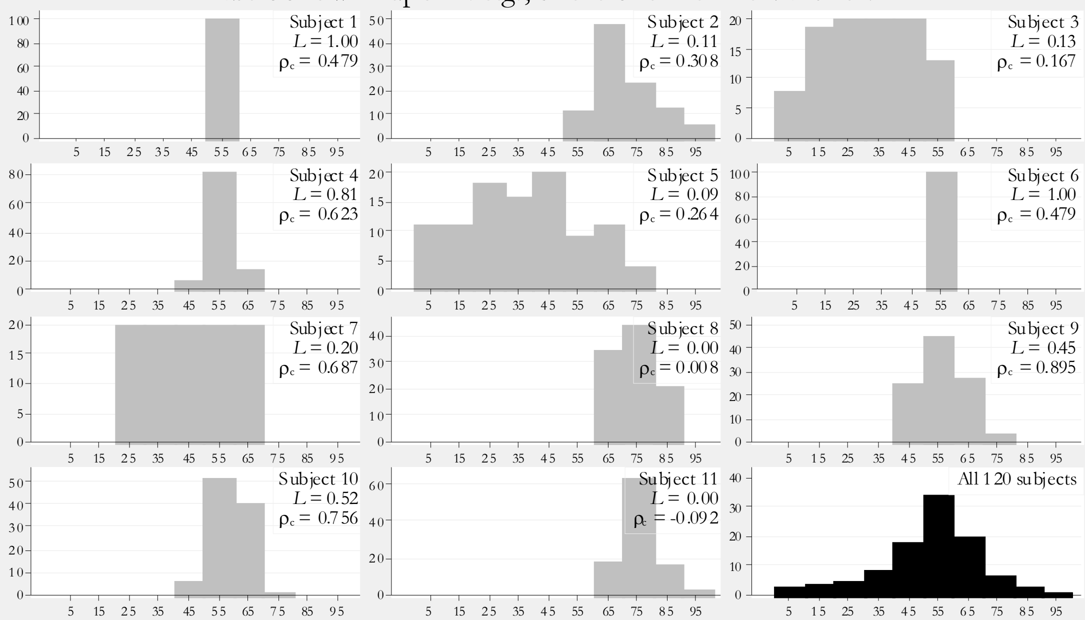


Figure 8: Elicited Beliefs For Remaining Lifetime

Correct answer is 59.1 years for men and 62.9 years for women

Average Literacy index $L = .34$ for men and .29 for women

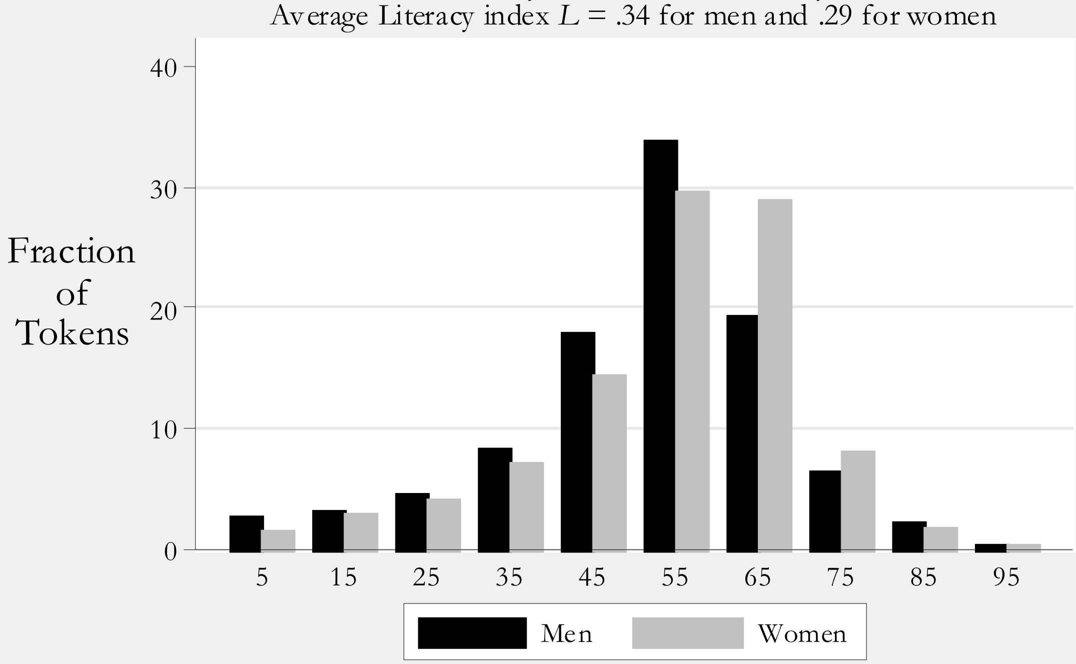


Figure 9: Beliefs of Breast Cancer Risk in Probability Format

Correct Answer is 7.8%

Average Literacy index L = 0.37 and average Concordance index ρ_c = 0.305

Labels show midpoint: e.g., 5% refers to the interval 0% to 9.9%

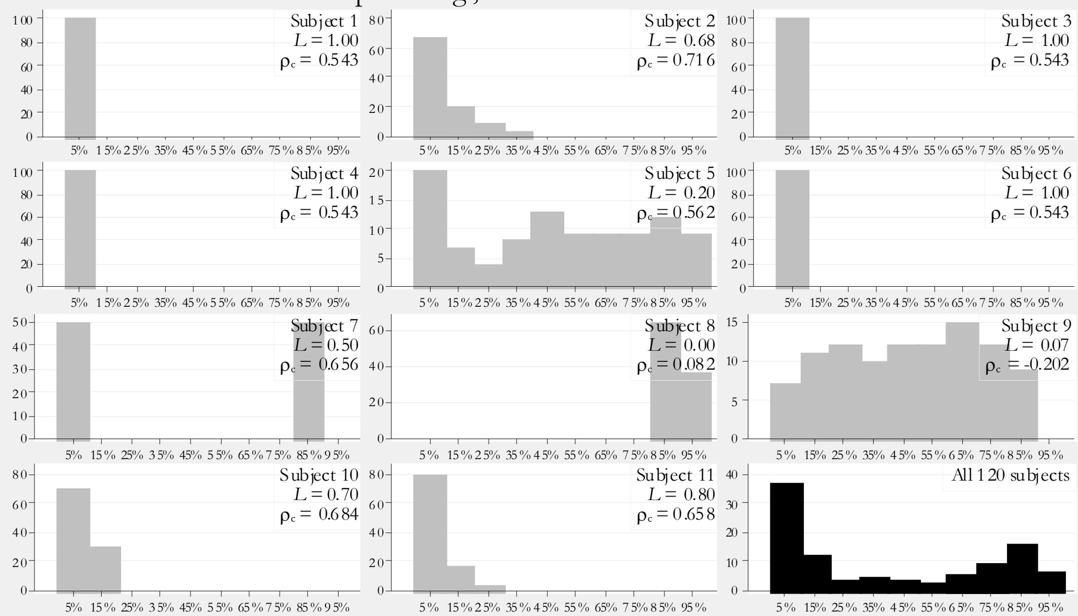


Figure 10: Elicited Beliefs For Breast Cancer Risk

Correct posterior probability is 7.8%

Average Literacy index L = .37 for Probability Format and .48 for Frequency Format

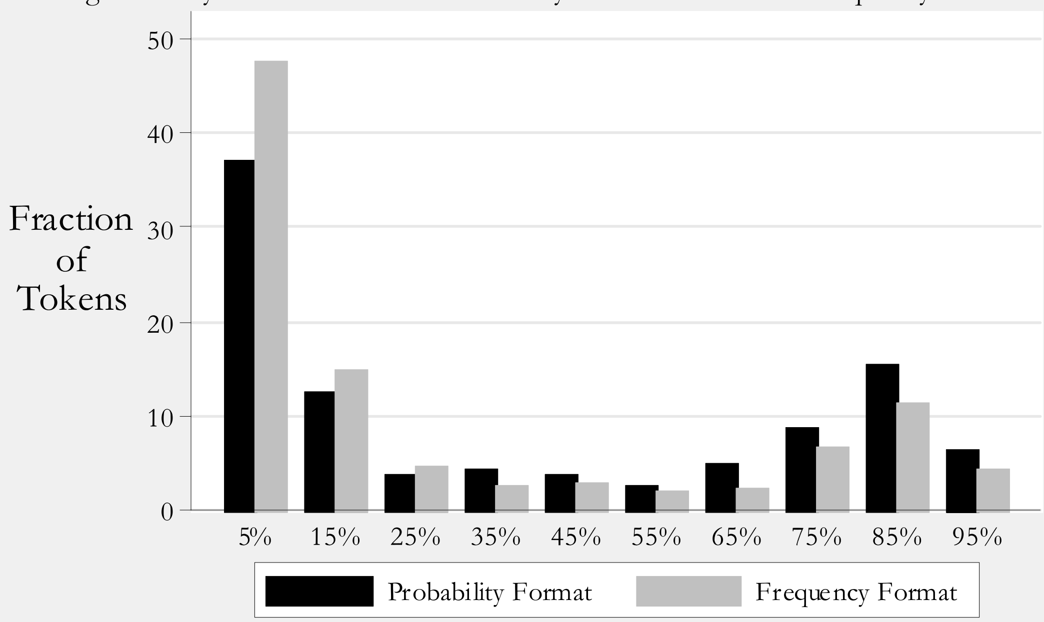


Figure 11: Beliefs of Credit Risk in Probability Format

Correct Answer is 67%
Average Literacy index L = 0.34 and average Concordance index ρ_c = 0.222
Labels show midpoint: e.g., 5% refers to the interval 0% to 9.9%

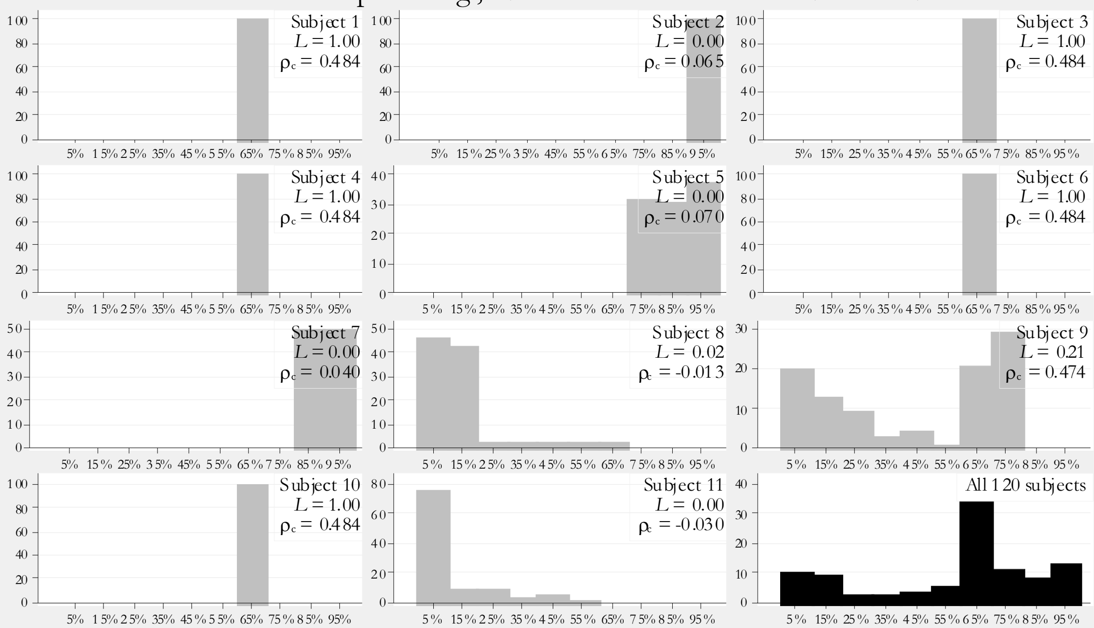


Figure 12: Elicited Beliefs For Credit Score Risk

Correct posterior probability is 67%
Average Literacy index L = .34 for Probability Format and .29 for Frequency Format

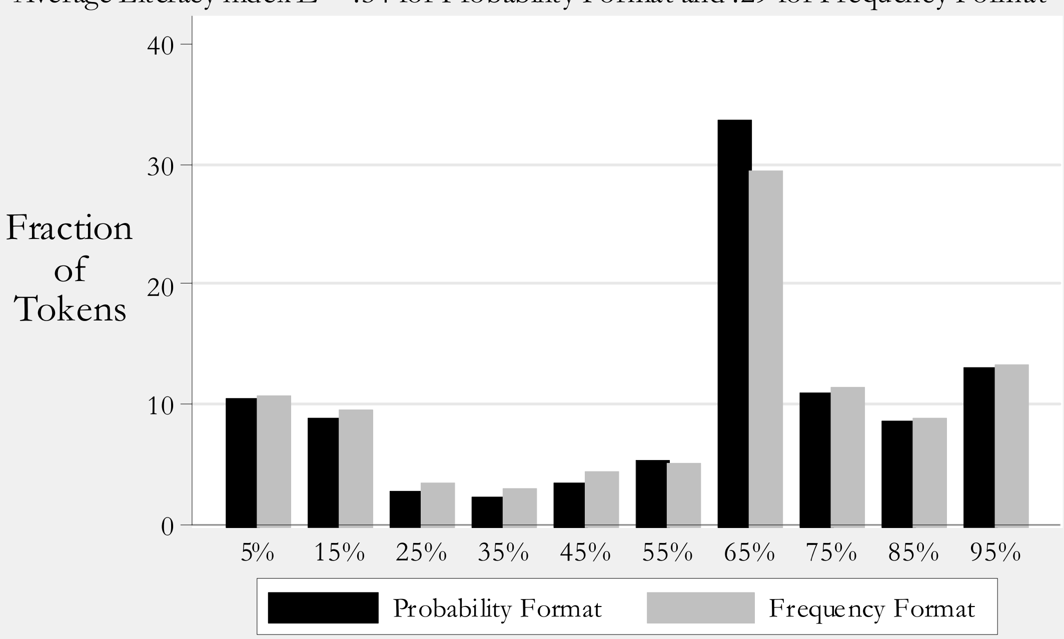


Figure 13: Literacy Indices for Economic Questions

Dashed line is the average

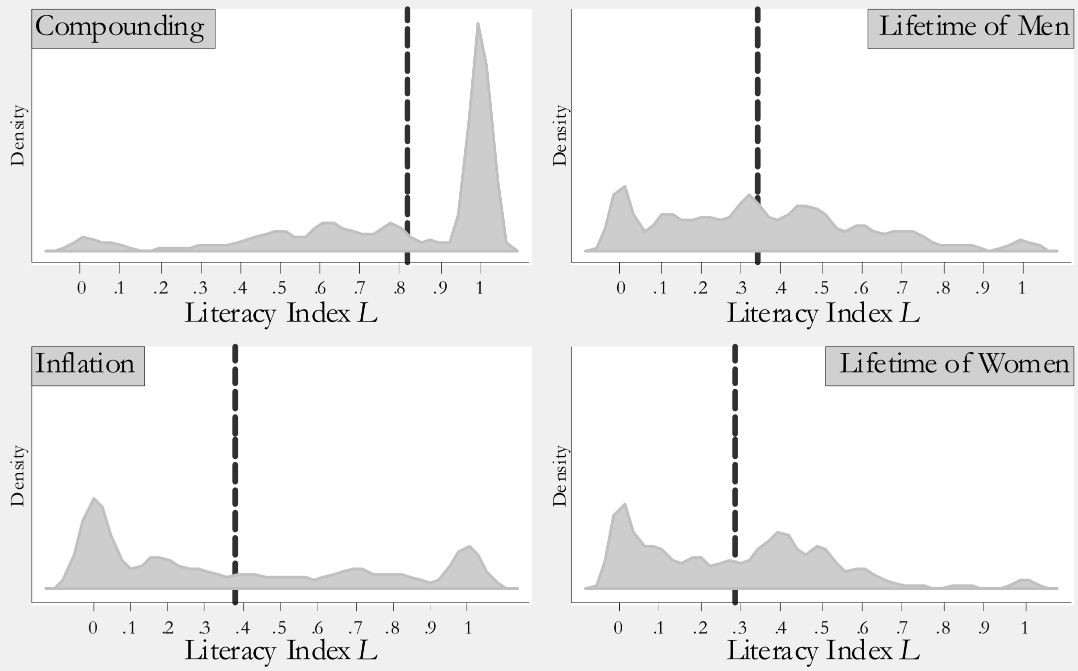


Figure 14: Literacy Indices for Statistical Questions

Dashed line is the average

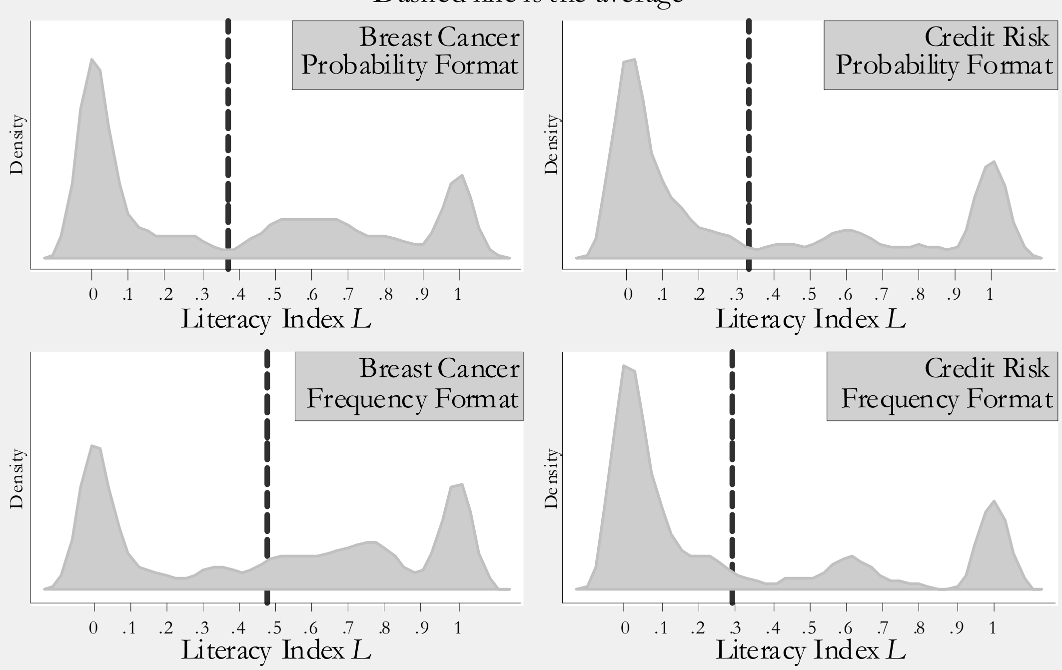


Table 1: Ordered Logit Model of Literacy Index

Variable	Estimate	Standard Error	*p*-value	95% Confidence Interval
		A. Effect on Probability of Being Illiterate		
Inflation	0.39	0.009	0.000	0.37 ↔ 0.40
Lifetime of men	0.41	0.008	0.000	0.39 ↔ 0.42
Lifetime of women	0.43	0.008	0.000	0.41 ↔ 0.44
Breast cancer, probability	0.36	0.008	0.000	0.35 ↔ 0.38
Breast cancer, frequency	-0.13	0.016	0.000	-0.16 ↔ -0.10
Credit risk, probability	0.45	0.009	0.000	0.43 ↔ 0.47
Credit risk, frequency	0.06	0.018	0.001	0.03 ↔ 0.09
Female	0.07	0.011	0.000	0.04 ↔ 0.09
Age, standardized	-0.05	0.007	0.000	-0.07 ↔ -0.04
Single	-0.15	0.018	0.000	-0.18 ↔ -0.11
White	0.03	0.015	0.074	-0.00 ↔ 0.06
Finance major	-0.06	0.010	0.000	-0.08 ↔ -0.04
Non-EU citizen	0.10	0.016	0.000	0.07 ↔ 0.13
Current smoker	0.02	0.013	0.241	-0.01 ↔ 0.04
Cognitive reflection test	-0.01	0.005	0.152	-0.02 ↔ 0.00
Berlin numeracy test	-0.10	0.005	0.000	-0.11 ↔ -0.09
		B. Effect on Probability of Being Semi-Literate		
Inflation	-0.11	0.004	0.000	-0.12 ↔ -0.10
Lifetime of men	-0.12	0.004	0.000	-0.12 ↔ -0.11
Lifetime of women	-0.12	0.004	0.000	-0.13 ↔ -0.12
Breast cancer, probability	-0.07	0.002	0.000	-0.07 ↔ -0.06
Breast cancer, frequency	0.02	0.002	0.000	0.02 ↔ 0.02
Credit risk, probability	-0.10	0.003	0.000	-0.11 ↔ -0.09
Credit risk, frequency	-0.01	0.005	0.002	-0.02 ↔ -0.05
Female	-0.01	0.002	0.000	-0.02 ↔ -0.01
Age, standardized	0.01	0.001	0.000	0.01 ↔ 0.01
Single	0.04	0.006	0.000	0.03 ↔ 0.05
White	-0.01	0.003	0.080	-0.01 ↔ 0.00
Finance major	0.01	0.002	0.000	0.01 ↔ 0.02
Non-EU citizen	-0.02	0.002	0.000	-0.02 ↔ -0.01
Current smoker	-0.00	0.003	0.256	-0.01 ↔ 0.00
Cognitive reflection test	0.00	0.001	0.153	-0.00 ↔ 0.00
Berlin numeracy test	0.02	0.001	0.000	0.02 ↔ 0.02
		C. Effect on Probability of Being Literate		
Inflation	-0.23	0.006	0.000	-0.29 ↔ -0.27
Lifetime of men	-0.29	0.006	0.000	-0.30 ↔ -0.28
Lifetime of women	-0.30	0.006	0.000	-0.31 ↔ -0.29
Breast cancer, probability	-0.30	0.007	0.000	-0.31 ↔ -0.28
Breast cancer, frequency	0.11	0.015	0.000	0.08 ↔ 0.14
Credit risk, probability	-0.35	0.007	0.000	-0.37 ↔ -0.34
Credit risk, frequency	-0.05	0.013	0.000	-0.07 ↔ -0.02

Female	-0.05	0.009	0.000	-0.07 ↔ -0.04
Age, standardized	0.04	0.005	0.000	0.03 ↔ 0.05
Single	0.11	0.013	0.000	0.08 ↔ 0.13
White	-0.02	0.012	0.073	-0.04 ↔ 0.00
Finance major	0.05	0.008	0.000	0.03 ↔ 0.06
Non-EU citizen	-0.08	0.014	0.000	-0.11 ↔ -0.06
Current smoker	-0.01	0.010	0.237	-0.03 ↔ 0.01
Cognitive reflection test	0.01	0.004	0.153	-0.00 ↔ 0.01
Berlin numeracy test	0.08	0.004	0.000	0.07 ↔ 0.09

Figure 15: Concordance Indices for Economic Questions

Dashed line is the average

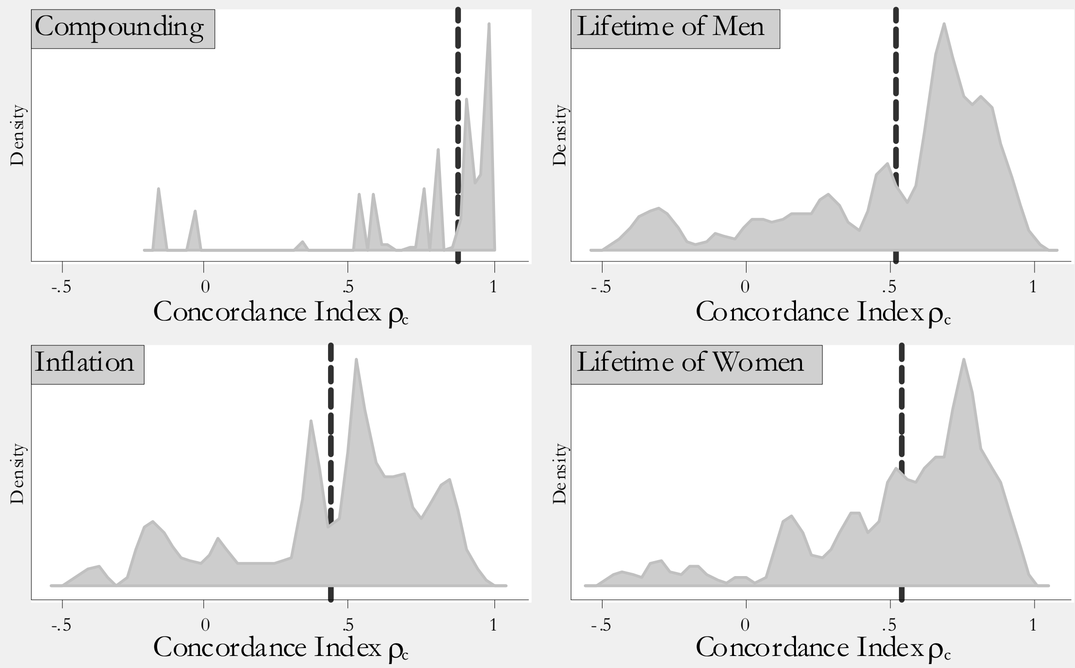


Figure 16: Concordance Indices for Statistical Questions

Dashed line is the average

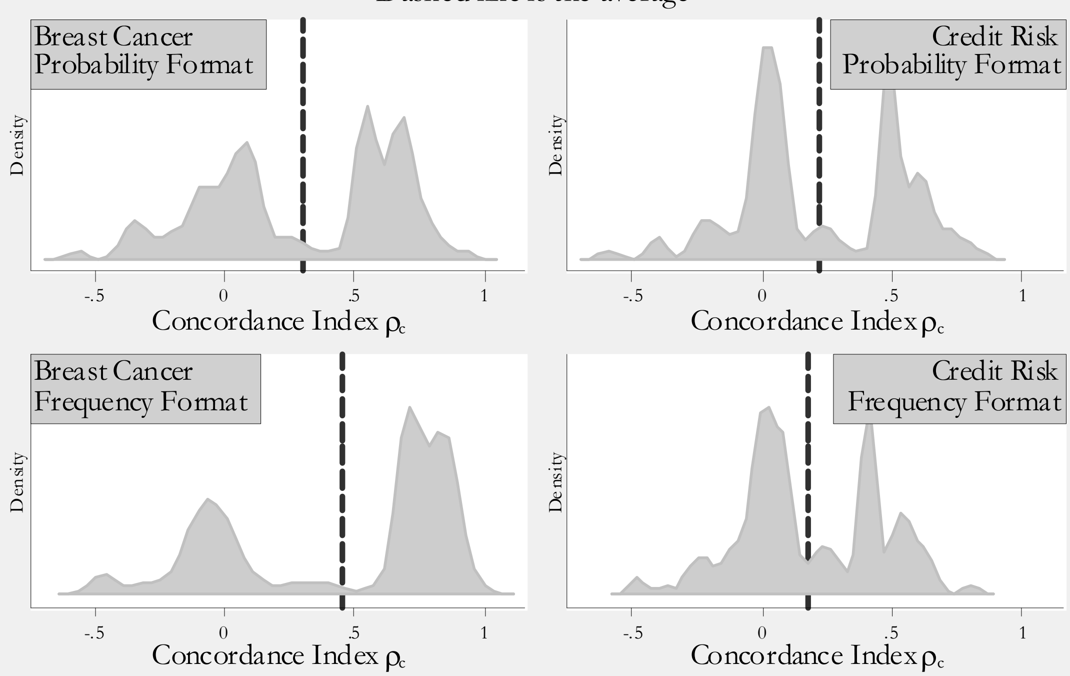


Figure 17: Literacy Rates and Subjective Expected Utility

Average of raw literacy rates given by index L and SEU-implied literacy rates if average belief is used

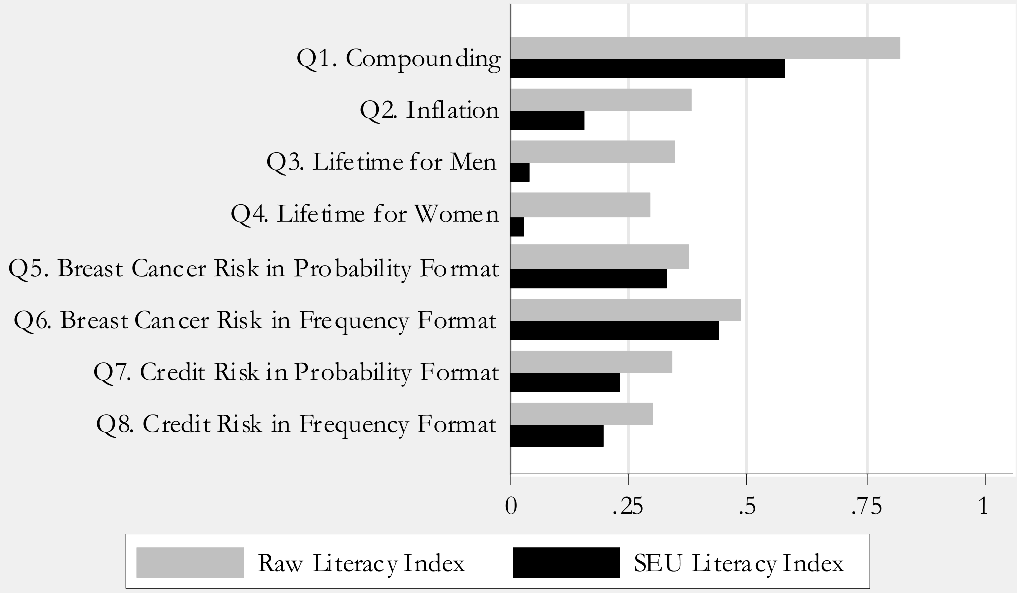


Figure 18: Distribution of p-values on ROCL Hypothesis Test Based on Individual Model of Behavior

Estimates obtained for 114 out of 120 subjects.

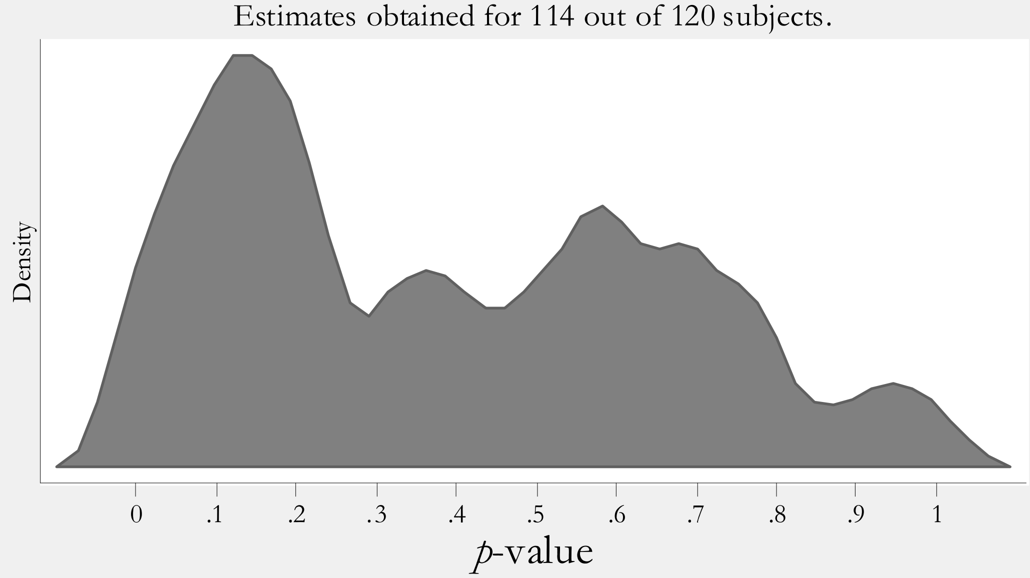
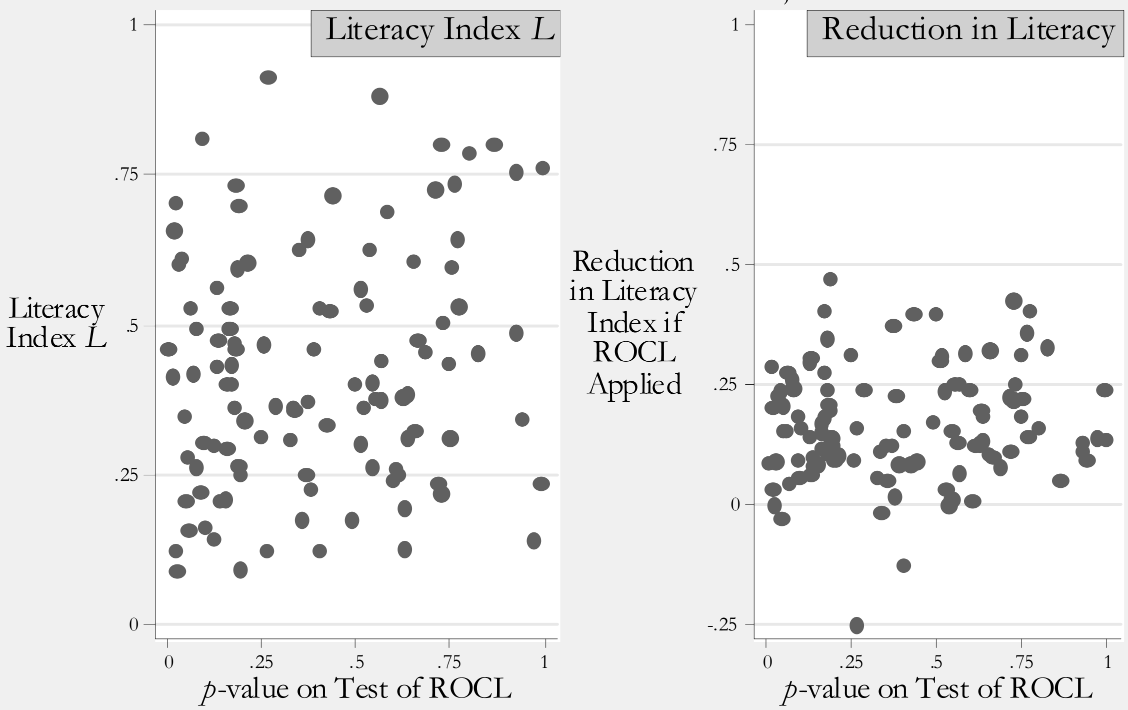


Figure 19: Literacy and ROCL

Data and estimates from 114 of 120 subjects



References

Allen, James, *Natural Language Understanding* (Redwood City, CA: Benjamin/Cummings Publishing, Second Edition, 1995).

Andersen, Steffen; Fountain, John; Harrison, Glenn W., and Rutström, E. Elisabet, "Estimating Subjective Probabilities," *Working Paper 2010-06*, Center for the Economic Analysis of Risk, Robinson College of Business, Georgia State University, 2010.

Armentier, Oliver; de Bruin, Wändi Bruine; Topa, Giorgio; van der Klaauw, Wilbert, and Zafar, Basit, "Inflation Expectations and Behavior: Do Survey Respondents Act on their Beliefs?" *Working Paper*, Federal Reserve Bank of New York, August 2011.

Basu, Laushik, and Foster, James E., "On Measuring Literacy," *Economic Journal*, 108, November 1998, 1733-1749.

Bateman, Hazel; Eckert, Christine; Geweke, John; Louviere, Jordan; Thorp, Susan, and Satchell, Stephen E., "Financial Competence and Expectations Formation: Evidence from Australia," *Economic Record*, 88, March 2012, 39-63.

Blackburn, McKinley; Harrison, Glenn W., and Rutström, E. Elisabet, "Statistical Bias Functions and Informative Hypothetical Surveys," *American Journal of Agricultural Economics*, 76(5), December 1994, 1084-1088.

Clark, Herbert H., *Arenas of Language Use* (Chicago: University of Chicago Press, 1992).

Cokely, Edward T.; Galesic, Mirta; Schulz, Eric; Ghazal, Saima and Garcia-Retamero, Rocio, "Measuring Risk Literacy: The Berlin Numeracy Test," *Judgment and Decision Making*, 7(1), January 2012, 25-47.

Delavande, Adeline; Giné, Xavier, and McKenzie, David, "Measuring Subjective Expectations in Developing Countries: A Critical Review and New Evidence," *Journal of Development Economics*, 94, 2011, 151-163.

Delevande, Adeline, and Kohler, Hans-Peter, "The Impact of HIV Testing on Subjective Expectations and Risky Behavior in Malawi," *Demography*, 49, 2012, 1011-1036.

El-Gamal, Mahmoud A., and Grether, David M., "Are people Bayesian? Uncovering behavioral strategies," *Journal of the American Statistical Association*, 90(432), 1995, 1137–1145.

Elder, Todd E., "The Predictive Validity of Subjective Mortality Expectations: Evidence from the Health and Retirement Study," *Demography*, 50, 2013 forthcoming.

Ellsberg, Daniel, "Risk, Ambiguity and the Savage Axioms," *Quarterly Journal of Economics*, 75(4), November 1961, 643-669.

Fehr, Ernst, and Tyran, Jean-Robert, "Does Money Illusion Matter?" *American Economic Review*, 91(5), December 2001, 1239-1262.

Fountain, John, and Harrison, Glenn W., "What Do Prediction Markets Predict?" *Applied Economics Letters*, 18, 2011, 267-272.

Frederick, Shane, "Cognitive Reflection and Decision Making," *Journal of Economic Perspectives*, 19(4), 2005, 25-42.

Ghirardato, Paolo; Maccheroni, Fabio, and Marinacci, Masimo, "Differentiating Ambiguity and Ambiguity Attitude," *Journal of Economic Theory*, 118, 2004, 133-173.

Gigerenzer, Gerd, "On Narrow Norms and Vague Heuristics: A Reply to Kahneman and Tversky," *Psychological Review*, 103, 1996, 592–596.

Gigerenzer, Gerd., and Hoffrage, Ulrich, "How to Improve Bayesian Reasoning Without Instructions: Frequency Formats," *Psychological Review*, 102(4), 1995, 684-704.

Grether, David M., "Testing Bayes Rule and the Representativeness Heuristic: Some Experimental Evidence," *Journal of Economic Behavior & Organization*, 17, 1992, 31-57.

Grice, Paul, *Studies in the Way of Words* (Cambridge, MA: Harvard University Press, 1989).

Harrison, Glenn W., "Theory and Misbehavior of First-Price Auctions," *American Economic Review*, 79, September 1989, 749-762.

Harrison, Glenn W., "Theory and Misbehavior of First-Price Auctions: Reply," *American Economic Review*, 82, December 1992, 1426-1443.

Harrison, Glenn W., "Expected Utility Theory and The Experimentalists," *Empirical Economics*, 19(2), 1994, 223-253.

Harrison, Glenn W., and Hirshleifer, Jack, "An Experimental Evaluation of Weakest-Link/Best-Shot Models of Public Goods," *Journal of Political Economy*, 97, February 1989, 201-225.

Harrison, Glenn W, Martínez-Correa, Jimmy and Swarthout, J. Todd, "Reduction of Compound Lotteries with Objective Probabilities: Theory and Evidence," *Working Paper 2012-05*, Center for the Economic Analysis of Risk, Robinson College of Business, Georgia State University, 2012.

Harrison, Glenn W, Martínez-Correa, Jimmy; Swarthout, J. Todd, and Ulm, Eric "Scoring Rules for Subjective Probability Distributions," *Working Paper 2012-10*, Center for the Economic Analysis of Risk, Robinson College of Business, Georgia State University, 2012.

Harrison, Glenn W., and Rutström, E. Elisabet, "Risk Aversion in the Laboratory," in J.C. Cox and G.W. Harrison (eds.), *Risk Aversion in Experiments* (Bingley, UK: Emerald, Research in

Experimental Economics, Volume 12, 2008).

Holt, Charles A., and Smith, Angela M., "An Update on Bayesian Updating," *Journal of Economic Behavior & Organization*, 69, 2009, 125-134.

Huston, Sandra J., "Measuring Financial Literacy," *Journal of Consumer Affairs*, 44(2), Summer 2010, 296-316.

Kadane, J. B. and Winkler, Robert L., "Separating Probability Elicitation from Utilities," *Journal of the American Statistical Association*, 83(402), 1988, 357-363.

Kahneman, Daniel, and Tversky, Amos, "On the Reality of Cognitive Illusions," *Psychological Review*, 103, 1996, 582–591.

Karni, Edi, "A Mechanism for Eliciting Probabilities," *Econometrica*, 77(2), March 2009, 603-606.

Klibanoff, Peter; Marinacci, Massimo, and Mukerji, Sujoy, "A Smooth Model of Decision Making under Ambiguity," *Econometrica*, 73(6), November 2005, 1849-1892.

Köszegi, Botond, and Rabin, Matthew, "Revealed Mistakes and Revealed Preferences," in A. Caplin and A. Schotter (eds.), *The Foundations of Positive and Normative Economics: A Handbook* (New York: Oxford University Press, 2008).

Lin, Lawrence I-Kuei, "A concordance correlation coefficient to evaluate reproducibility," *Biometrics*, 45, 1989, 255–268.

Lin, Lawrence I-Kuei, "A note on the concordance correlation coefficient," *Biometrics*, 56, 2000, 324–325.

Lusardi, Annamaria, and Mitchell, Olivia S., "Baby Boomer Retirement Security: The Roles of Planning, Financial Literacy, and Housing Wealth," *Journal of Monetary Economics*, 54, 2007, 205-224.

Lusardi, Annamaria, and Mitchell, Olivia S., "Planning and Financial Literacy: How do Women Fare?" *American Economic Review (Papers & Proceedings)*, 98(2), 2008, 413-417.

Lusardi, Annamaria, and Mitchell, Olivia S., *Financial Literacy: Implications for Retirement Security and the Financial Marketplace* (New York: Oxford University Press, 2012).

Lusardi, Annamaria; Mitchell, Olivia S., and Curto, Vilsa, "Financial Literacy Among the Young," *Journal of Consumer Affairs*, 44(2), Summer 2010, 358-380.

Manski, Charles F., "Interpreting the Predictions of Prediction Markets," *Economics Letters*, 91, 2006, 425–9.

Matheson, James E., and Winkler, Robert L, "Scoring Rules for Continuous Probability

Distributions," *Management Science*, 22(10), June 1976, 1087-1096.

Nickerson, Carol A. E., "A Note on 'A Concordance Correlation Coefficient to Evaluate Reproducibility'," *Biometrics*, 53, December 1997, 1503-1507.

Perozek, Maria, "Using Subjective Expectations to Forecast Longevity: Do Survey Respondents Know Something We Don't Know?" *Demography*, 45(1), February 2008, 95-113.

Remund, David L., "Financial Literacy Explicated: The Case for a Clearer Definition in an Increasingly Complex Economy," *Journal of Consumer Affairs*, 44(2), Summer 2010, 276-295.

Savage, Leonard J., "Elicitation of Personal Probabilities and Expectations," *Journal of American Statistical Association*, 66, December 1971, 783-801.

Savage, Leonard J., *The Foundations of Statistics* (New York: Dover Publications, 1972; Second Edition).

Shefrin, Hersh, *A Behavioral Approach to Asset Pricing* (New York: Academic Press, Second Edition, 2008).

Smith, V. Kerry; Taylor, Donald H., and Sloan, Frank A., "Longevity and Expectations: Can People Predict Their Own Demise?" *American Economic Review*, 91, September 2001, 1126-1134.

van Rooij, Maarten; Lusardi, Annamaria and Alessie, Rob, "Financial Literacy and Stock Market Participation," *Journal of Financial Economics*, 101(2), August 2011, 449–472.

Winkler, Robert L., and Murphy, Allan H., "Nonlinear Utility and the Probability Score," *Journal of Applied Meteorology*, 9, February 1970, 143-148.

Appendix A: Experimental Instructions (NOT FOR PUBLICATION)

Text in square brackets is for the experimenter only.

A. General Instructions

Introduction

You are now participating in a decision-making experiment. Based on your decisions in this experiment, you can earn money that will be paid to you in cash today. It is important that you understand all instructions before making your choices in this experiment.

Please turn to silent, and put away, your mobile phone, laptop computer, or any other device you may have brought with you. Please do not talk with others seated nearby for the duration of the experiment. If at any point you have a question, please raise your hand and we will answer you as soon as possible.

The experiment consists of a demographic survey and two decision-making tasks. You have already earned £5 for agreeing to participate in the experiment, which will be paid in cash at the end of the session. In addition to this show-up fee, you may earn considerably more from your choices in the decision-making tasks. These tasks and the potential earnings from them will be explained in detail as we proceed through the session.

Before we begin the experiment, we will give you an informed consent form. This form explains your rights as a participant in the experiment, what the experiment is about and how we make payments to you. [Give the informed consent form to subjects and read it out loud.]

We will begin the experiment by asking you to answer some demographic questions. Your answers to those questions will not influence your payoffs. [Give the questionnaire to subjects.]

We will now continue with the first decision-making task. You will be given written instructions, but make all decisions on the computer in front of you. We will distribute the instructions and then read it out loud. Please remain silent during the experiment, and simply raise your hand if you have any question so that an experimenter will come to you.

[Give the first set of instructions to subjects and read it out loud.]
[Determine earnings for each subject.]

We will now continue with the second decision-making task. You will again be given written instructions and make all decisions on the computer in front of you. We will distribute the instructions and then read it out loud. Please remain silent during the experiment, and simply raise your hand if you have any question so that an experimenter will come to you.

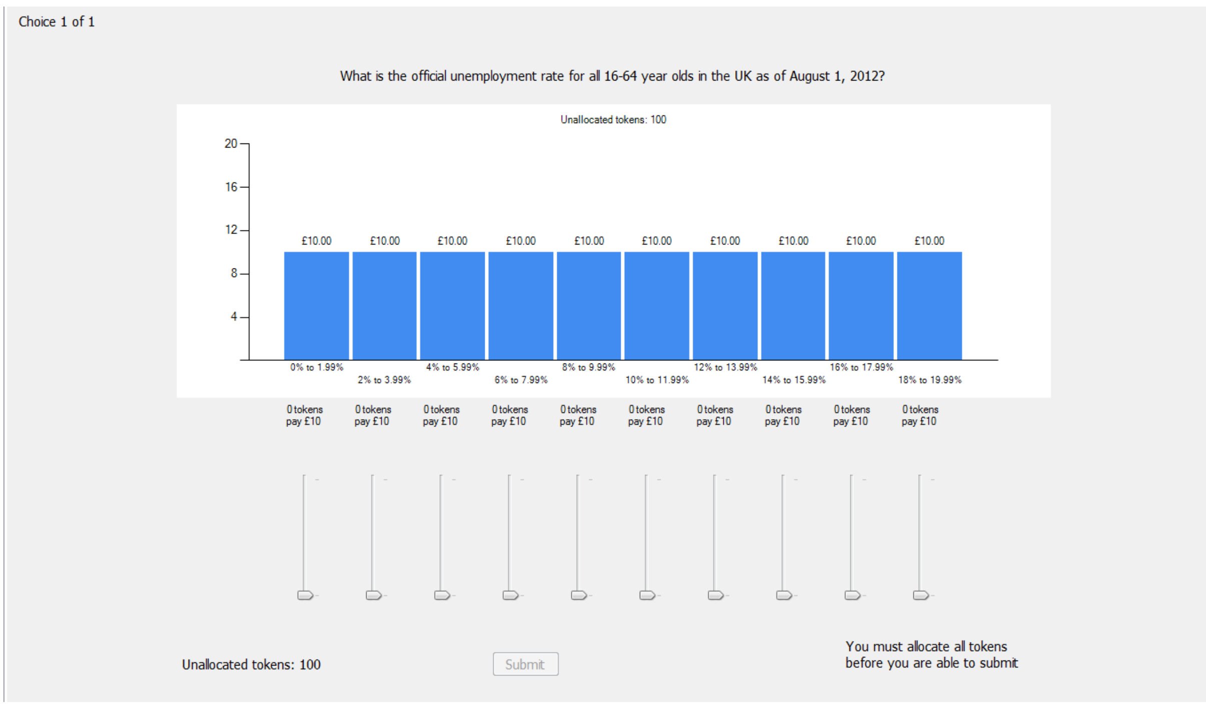
[Give the second set of instructions to subjects and read it out loud.]
[Determine earnings for each subject.]

B. Belief Elicitation Instructions

Your Beliefs

This is a task where you will be paid according to how accurate your beliefs are about certain things. You will be presented with 8 questions of the type we will explain below. You will actually get the chance to play one question presented to you, so you should think carefully about your answer to each question.

Here is an example of what the computer display of such a question might look like.

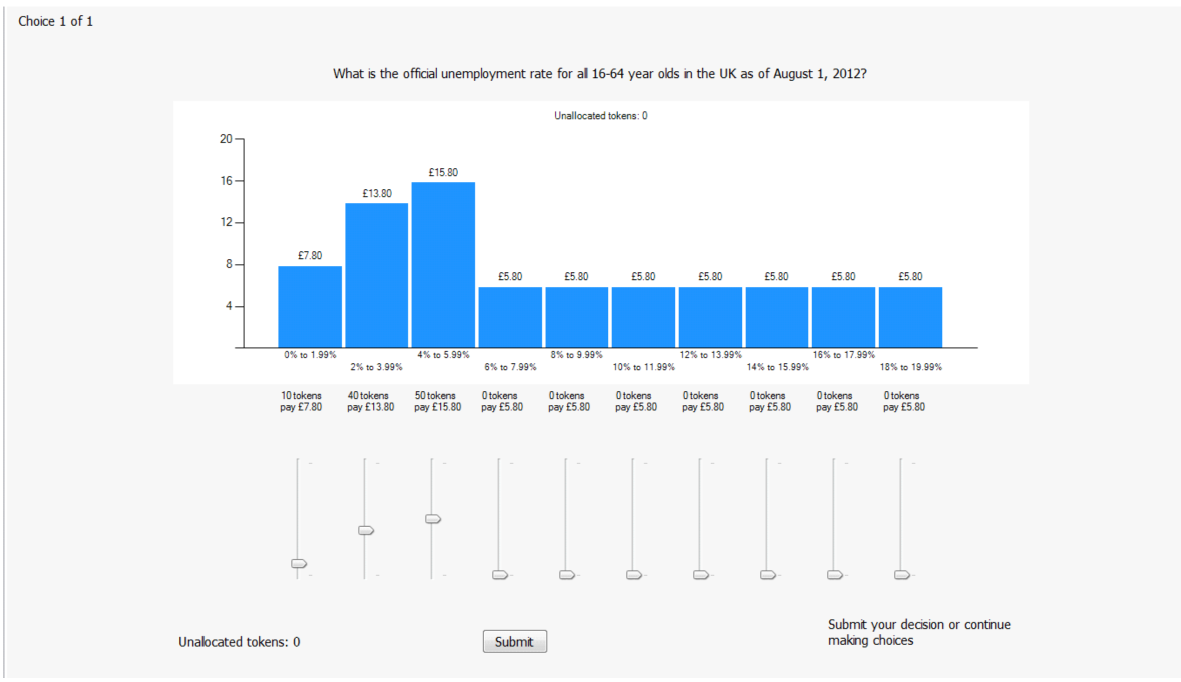


The display on your computer will be larger and easier to read. You have 10 sliders to adjust, shown at the bottom of the screen, and you have 100 tokens to allocate. Each slider allows you to allocate tokens to reflect your belief about the answer to this question. You must allocate all 100 tokens in order to submit your decision, and in this example we start with 10 tokens allocated to each slider. The payoffs shown on the screen only apply when you allocate all 100 tokens. As you allocate tokens, by adjusting sliders, the payoffs displayed on the screen will change. Your earnings are based on the payoffs that are displayed after you have allocated all 100 tokens.

You can earn up to £20 in this task.

Where you position each slider depends on your beliefs about the correct answer to the question. In the above example the tokens you allocate to each bar will naturally reflect your beliefs about the official unemployment rate for all 16-64 years old in the UK as of August 1, 2012. The first bar here corresponds to your belief that the unemployment rate is between 0% and 1.99%. The second bar corresponds to your belief that the unemployment rate is between 2% and 3.99%, and so on. Each bar here shows the amount of money you earn if the official unemployment rate is in the interval shown under the bar.

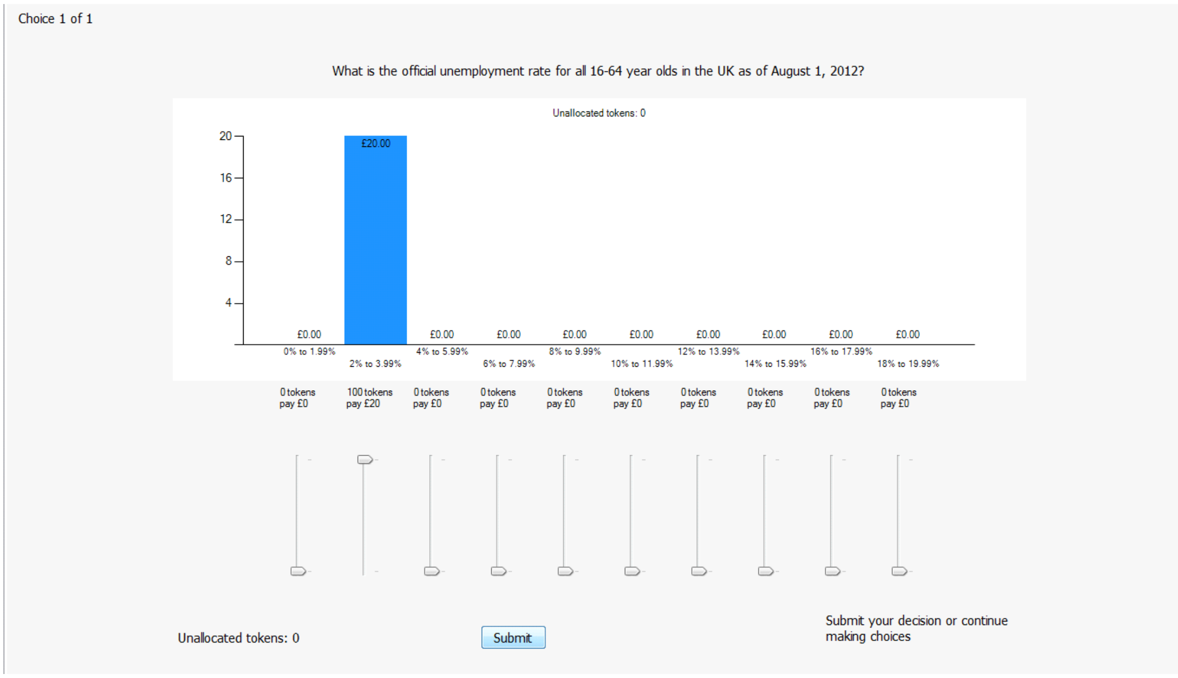
To illustrate how you use these sliders, suppose you think there is a fair chance the unemployment rate is just under 5%. Then you might allocate the 100 tokens in the following way: 50 tokens to the interval 4% to 5.99%, 40 tokens to the interval 2% to 3.99%, and 10 tokens to the interval 0% to 1.99%. So you can see in the picture below that if indeed the unemployment rate is between 4% to 5.99% you would earn £15.80. You would then earn less than £15.80 for any other outcome. You would earn £13.80 if the unemployment rate is between 2% and 3.99%, £7.80 if it is between 0% and 1.99%, and for any other unemployment rate you would earn £5.80.



You can adjust the allocation as much as you want to best reflect your personal beliefs about the unemployment rate.

Your earnings depend on your reported beliefs and, of course, the true answer. For instance, suppose you allocated your tokens as in the figure shown above. The true unemployment is 7.8%. So if you had reported the beliefs shown above, you would have earned £5.80.

Suppose you had put all of your eggs in one basket, and for example allocated 100 tokens to the interval corresponding to 2.5%. Then you would have faced the earnings outcomes shown below.



Note the "good news" and "bad news" here. If the unemployment rate is indeed between 2% and 3.99%, you earn the maximum payoff, shown here as £20. But the true unemployment rate is 7.8%, so you would have earned nothing in this task.

It is up to you to balance the strength of your personal beliefs with the risk of them being wrong. There are three important points for you to keep in mind when making your decisions:

- **Your belief about the correct answer to each question is a personal judgment that depends on the information you have about the different events.**

- **Depending on your choices and the correct answer you can earn up to £20.**

- **Your choices might also depend on your willingness to take risks or to gamble.**

The decisions you make are a matter of personal choice. Please work silently, and make your choices by thinking carefully about the questions you are presented with.

When you are happy with your decisions, you should click on the **Submit** button and confirm your choices. When everyone is finished we will reveal the right answer to each of the 8 questions. Then an experimenter will come to you and ask you to roll a 10-sided die until a number between 1 and 8 comes up to determine which question will be played out. The experimenter will record your earnings according to the correct answer and the choices you made.

All payoffs are in cash, and are in addition to the £5 show-up fee that you receive just for being here.

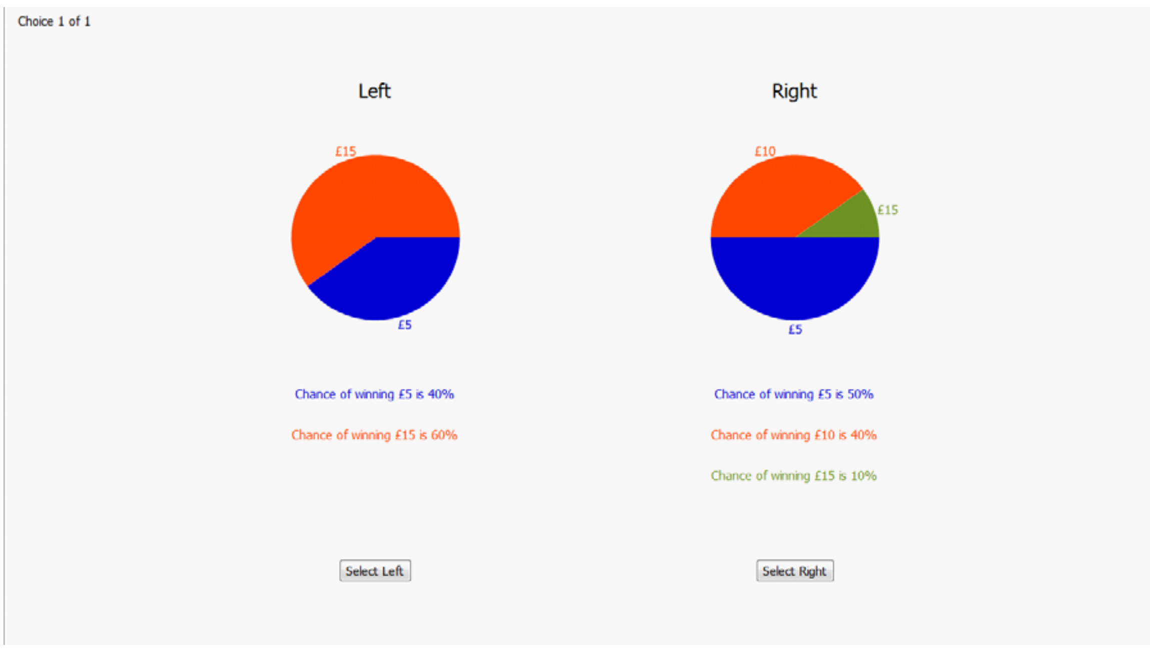
Are there any questions?

C. Risky Lotteries Task

Choices over Prospects

This is a task where you will choose between prospects with varying prizes and chances of winning. You will be presented with 40 pairs of prospects. For each pair of prospects, you should choose the prospect you prefer to play. You will actually get the chance to play **one** of the prospects you choose, and you will be paid according to the outcome of that prospect, so you should think carefully about which prospect you prefer.

Here is an example of what the computer display of such a pair of prospects might look like.



The outcome of the prospects will be determined by the draw of a random number between 1 and 100. Each number between, and including, 1 and 100 is equally likely to occur. In fact, you will be able to draw the number yourself using two 10-sided dice.

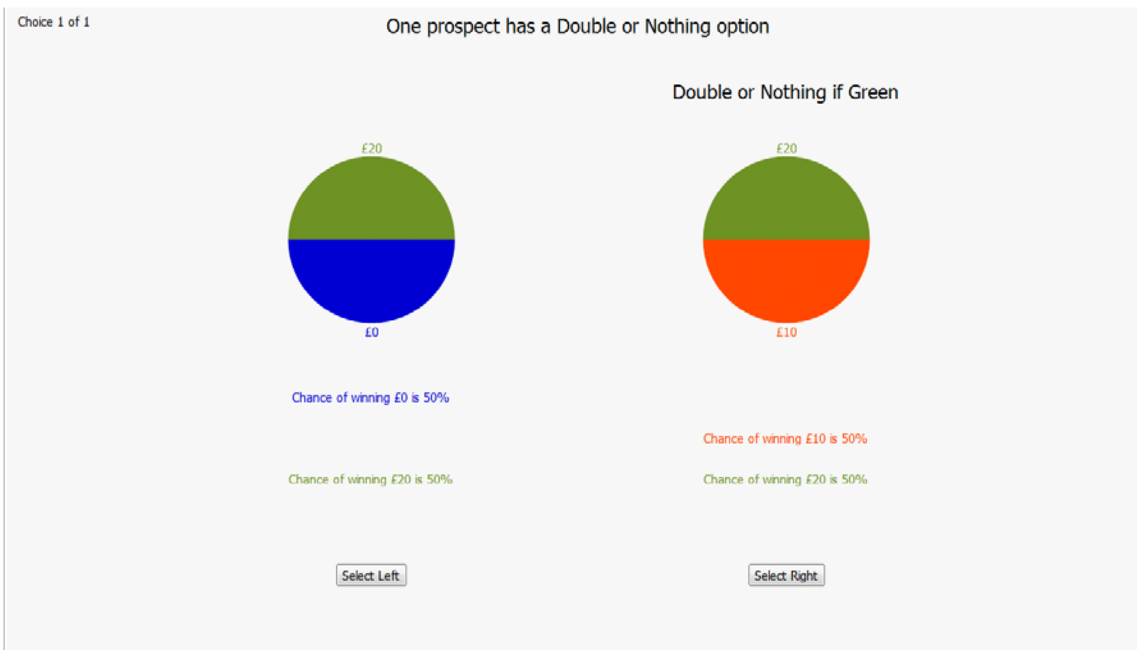
In the above example the left prospect pays five pounds (£5) if the number drawn is between 1 and 40, and pays fifteen pounds (£15) if the number is between 41 and 100. The blue color in the pie chart corresponds to 40% of the area and illustrates the chances that the number drawn will be between 1 and 40 and your prize will be £5. The orange area in the pie chart corresponds to 60% of the area and illustrates the chances that the number drawn will be between 41 and 100 and your prize will be £15.

Now look at the pie in the chart on the right. It pays five pounds (£5) if the number drawn is between 1 and 50, ten pounds (£10) if the number is between 51 and 90, and fifteen pounds (£15) if the number is between 91 and 100. As with the prospect on the left, the pie slices represent the fraction of the possible numbers which yield each payoff. For example, the size of the £15 pie slice is

10% of the total pie.

Each pair of prospects is shown on a separate screen on the computer. On each screen, you should indicate which prospect you prefer to play by clicking on one of the buttons beneath the prospects.

You could also get a pair of prospects in which one of the prospects will give you the chance to play "Double or Nothing." For instance, the right prospect in the following screen image pays "Double or Nothing" if the Green area is selected, which happens if the number drawn is between 51 and 100. The right pie chart indicates that if the number is between 1 and 50 you get £10. However, if the number is between 51 and 100 a coin will be tossed to determine if you get double the amount. If it comes up Heads you get £40, otherwise you get nothing. The prizes listed underneath each pie refer to the amounts <u>before</u> any "Double or Nothing" coin toss.



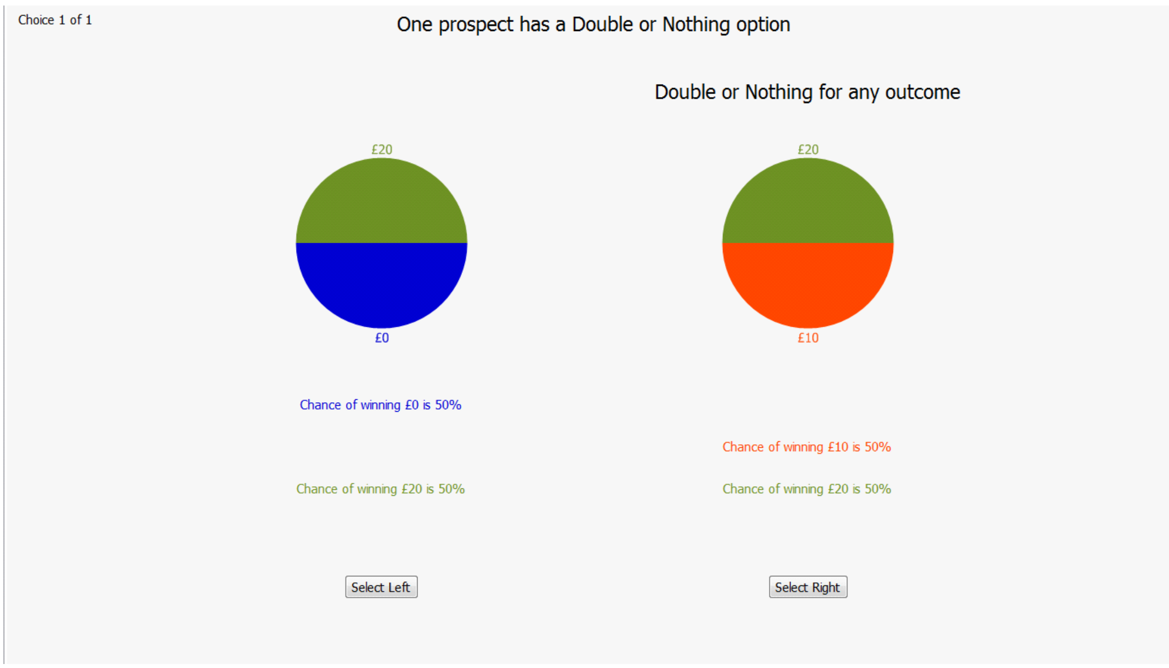
After you have worked through all of the 40 pairs of prospects, raise your hand and an experimenter will come over. You will then roll two 10-sided dice until a number between 1 and 40 comes up to determine which pair of prospects will be played out. Since there is a chance that any of your 40 choices could be played out for real, you should approach each pair of prospects as if it is the one that you will play out. Finally, you will roll the two ten-sided dice to determine the outcome of the prospect you chose, and if necessary you will then toss a coin to determine if you get "Double or Nothing."

For instance, suppose you picked the prospect on the left in the last example. If the random number was 37, you would win £0; if it was 93, you would get £20.

If you picked the prospect on the right and drew the number 37, you would get £10; if it was

93, you would have to toss a coin to determine if you get "Double or Nothing." If the coin comes up Heads then you get £40. However, if it comes up Tails you get nothing from your chosen prospect.

It is also possible that you will be given a prospect in which there is a "Double or Nothing" option no matter what the outcome of the random number. The screen image below illustrates this possibility.



Therefore, your payoff is determined by four things:

- by which prospect you selected, the left or the right, for each of these 40 pairs;
- by which prospect pair is chosen to be played out in the series of 40 such pairs using the two 10-sided dice;
- by the outcome of that prospect when you roll the two 10-sided dice; and
- by the outcome of a coin toss if the chosen prospect outcome is of the "Double or Nothing" type.

Which prospects you prefer is a matter of personal taste. The people next to you may be presented with different prospects, and may have different preferences, so their responses should not matter to you. Please work silently, and make your choices by thinking carefully about each prospect. All payoffs are in cash, and are in addition to the £5 show-up fee that you receive just for being here.

D. Demographic and Other Hypothetical Questions

ID ______________

In this survey most of the questions asked are descriptive. We will not be grading your answers and your responses are completely confidential. Please think carefully about each question and give your best answers.

1. What is your age? ____________ years

2. What is your sex? (Circle one number.)

01 Male
02 Female

3. Which of the following categories best describes you? (Circle one number.)

01 White
02 Mixed
03 Asian or Asian British
04 Chinese or other ethnic group
05 Prefer not to say

4. What is your main field of study? (Circle one number.)

01 Accounting
02 Economics
03 Finance
04 Business Administration, other than Accounting, Economics, or Finance
05 Education
06 Engineering
07 Health and Medicine
08 Biological and Biomedical Sciences
09 Math, Computer Sciences, or Physical Sciences
10 Social Sciences or History
11 Law
12 Psychology
13 Modern Languages and Cultures
14 Other Fields

5. What is your year of studies? (Circle one number.)

01 First year
02 Second year
03 Third year
04 Masters

05 Doctoral

6. What is the highest level of education you expect to complete? (Circle one number)

01 Bachelor's degree
02 Master's degree
03 Doctoral degree
04 Professional qualification

7. As a percentage, what is your current average mark if you are doing a Bachelor's degree, or what was it when you did a Bachelor's degree? This mark should refer to all your years of study for this degree, not just the current year. Please pick one by rounding up or down to the nearest number:

01 Above 70%
02 Between 60 – 69%
03 Between 50 – 59%
04 Between 40 – 49%
05 Less than 40%
06 Have not taken courses for which grades are given.

8. What is your citizenship status?

01 British Citizen
02 EU Citizen (non-British Citizen)
03 Non-EU Citizen

9. Are you currently:

01 Single and never married?
02 Married?
03 Separated, divorced or widowed?

10. How many people live in your household? Include yourself, your spouse and any dependents. Do not include your parents or roommates unless you claim them as dependents.

____

11. Please circle the category below that describes the total amount of income before tax earned in the calendar year 2007 by the people in your household (as "household" is defined in question 10). [Consider all forms of income, including salaries, tips, interest and dividend payments, scholarship support, student loans, parental support, social security, alimony, and child support, and others.]

01 Less than £10,000
02 £10,000 – £19,999
03 £20,000 – £29,999
04 £30,000 – £49,999
05 Over £50,000

12. Please circle the category below that describes the total amount of income before tax earned in the calendar year 2007 by your parents. [Consider all forms of income, including salaries, tips, interest and dividend payments, social security, alimony, and child support, and others.]

01 Less than £10,000
02 £10,000 – £19,999
03 £20,000 – £29,999
04 £30,000 – £49,999
05 Over £50,000
06 Don't Know

13. Do you currently smoke cigarettes? (Circle one number.)

00 No
01 Yes

If yes, approximately how much do you smoke in one day? ________ cigarettes.

14. A bat and a ball cost £1.10 in total. The bat costs £1.00 more than the ball. How much does the ball cost? £________

15. If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? ________ minutes

16. In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake? ________ days

17. Imagine we are throwing a five-sided die 50 times. On average, out of these 50 throws how many times would this five-sided die show an odd number (1, 3 or 5)? ________ out of 50 throws.

18. Out of 1,000 people in a small town 500 are members of a choir. Out of these 500 members in the choir 100 are men. Out of the 500 inhabitants that are not in the choir 300 are men. What is the probability that a randomly drawn man is a member of the choir? (Please indicate the probability in percent). ________%

19. Imagine we are throwing a loaded die (6 sides). The probability that the die shows a 6 is twice as high as the probability of each of the other numbers. On average, out of these 70 throws, how many times would the die show the number 6? ________ out of 70 throws.

20. In a forest 20% of mushrooms are red, 50% brown and 30% white. A red mushroom is poisonous with a probability of 20%. A mushroom that is not red is poisonous with a probability of 5%. What is the probability that a poisonous mushroom in the forest is red? ________%

E. Lottery Parameters

Lottery parameters directly from Harrison, Martínez-Correa and Swarthout [2012]. For our experiments all prizes were divided by 4 and presented as British pounds.

Table A1: Simple Lotteries vs. Compound Lotteries (Pairs 1-15)

								Compound Lottery								
					Simple Lottery			Initial Lottery			Initial Lottery					
		Final Prizes			Probabilities			Prizes			Probabilities			"Double or Nothing"	EV	EV
Pair	Context	Low	Middle	High	Low	Middle	High	Low	Middle	High	Low	Middle	High	option	Simple	Compound
1	1	$0	$10	$20	0.5	0.5	0	$0	$10	$20	0.5	0.5	0	DON if middle	$5.00	$5.00
2	1	$0	$10	$20	0	1	0	$0	$10	$20	0.5	0.5	0	DON if middle	$10.00	$5.00
3	2	$0	$10	$35	0	1	0	$0	$5	$17.50	0	0.5	0.5	DON for any outcome	$10.00	$11.25
13	2	$0	$10	$35	0.25	0.75	0	$0	$17.50	$35	0.5	0.5	0	DON if middle	$7.50	$8.75
4	3	$0	$10	$70	0.25	0.75	0	$0	$35	$70	0	1	0	DON for any outcome	$7.50	$35.00
5	3	$0	$10	$70	0	1	0	$0	$35	$70	0	1	0	DON for any outcome	$10.00	$35.00
6	4	$0	$20	$35	0	1	0	$0	$10	$35	0	0.5	0.5	DON if middle	$20.00	$22.50
14	4	$0	$20	$35	0	0.75	0.25	$0	$17.50	$35	0	0.5	0.5	DON if middle	$23.75	$8.75
7	5	$0	$20	$70	0	0.5	0.5	$0	$35	$70	0	0.5	0.5	DON if middle	$45.00	$52.50
9	5	$0	$20	$70	0.5	0	0.5	$0	$20	$35	0	0.5	0.5	DON if high	$35.00	$27.50
11	5	$0	$20	$70	0	1	0	$0	$20	$35	0	0.5	0.5	DON if high	$20.00	$27.50
15	5	$0	$20	$70	0	0.75	0.25	$0	$35	$70	0	0.5	0.5	DON if middle	$32.50	$52.50
8	6	$0	$35	$70	0	1	0	$0	$35	$70	0	0.5	0.5	DON if middle	$35.00	$52.50
10	6	$0	$35	$70	0	0.75	0.25	$0	$35	$70	0	1	0	DON for any outcome	$43.75	$35.00
12	6	$0	$35	$70	0	0.75	0.25	$0	$35	$70	0	0.5	0.5	DON if middle	$43.75	$52.50

Table A2: Simple Lotteries vs. Actuarially-Equivalent Lotteries (Pairs 16-30)

Pair	Context	Final Prizes Low	Final Prizes Middle	Final Prizes High	Simple Lottery Probabilities Low	Simple Lottery Probabilities Middle	Simple Lottery Probabilities High	Actuarially-Equivalent Lottery Probabilities Low	Actuarially-Equivalent Lottery Probabilities Middle	Actuarially-Equivalent Lottery Probabilities High	EV Simple	EV Actuarially-Equivalent
16	1	$0	$10	$20	0.5	0.5	0	0.75	0	0.25	$5.00	$5.00
17	1	$0	$10	$20	0	1	0	0.75	0	0.25	$10.00	$5.00
18	2	$0	$10	$35	0	1	0	0.5	0.25	0.25	$10.00	$11.25
19	2	$0	$10	$35	0.25	0.75	0	0.75	0	0.25	$7.50	$8.75
20	3	$0	$10	$70	0.25	0.75	0	0.5	0	0.5	$7.50	$35.00
21	3	$0	$10	$70	0	1	0	0.5	0	0.5	$10.00	$35.00
22	4	$0	$20	$35	0	1	0	0.25	0.25	0.5	$20.00	$22.50
23	4	$0	$20	$35	0	0.75	0.25	0.25	0	0.75	$23.75	$8.75
24	5	$0	$20	$70	0	0.5	0.5	0.25	0	0.75	$45.00	$52.50
25	5	$0	$20	$70	0.5	0	0.5	0.25	0.5	0.25	$35.00	$27.50
26	5	$0	$20	$70	0	1	0	0.25	0.5	0.25	$20.00	$27.50
27	5	$0	$20	$70	0	0.75	0.25	0.25	0	0.75	$32.50	$52.50
28	6	$0	$35	$70	0	1	0	0.25	0	0.75	$35.00	$52.50
29	6	$0	$35	$70	0	0.75	0.25	0.5	0	0.5	$43.75	$35.00
30	6	$0	$35	$70	0	0.75	0.25	0.25	0	0.75	$43.75	$52.50

Table A3: Actuarially-Equivalent Lotteries vs. Compound Lotteries (Pairs 31-40)

								Compound Lottery								
	Final Prizes				Actuarially-Equivalent Lottery Probabilities			Initial Lottery Prizes			Initial Lottery Probabilities			"Double or Nothing"	EV Actuarially-	EV Compound
Pair	Context	Low	Middle	High	Low	Middle	High	Low	Middle	High	Low	Middle	High	option	Equivalent	
1	1	$0	$10	$20	0.75	0	0.25	$0	$10	$20	0.5	0.5	0	DON if middle	$5.00	$5.00
3	2	$0	$10	$35	0.5	0.25	0.25	$0	$5	$17.50	0	0.5	0.5	DON for any outcome	$11.25	$11.25
13	2	$0	$10	$35	0.75	0	0.25	$0	$17.50	$35	0.5	0.5	0	DON if middle	$8.75	$8.75
4	3	$0	$10	$70	0.5	0	0.5	$0	$35	$70	0	1	0	DON for any outcome	$35.00	$35.00
6	4	$0	$20	$35	0.25	0.25	0.5	$0	$10	$35	0	0.5	0.5	DON if middle	$22.50	$22.50
14	4	$0	$20	$35	0.25	0	0.75	$0	$17.50	$35	0	0.5	0.5	DON if middle	$8.75	$8.75
7	5	$0	$20	$70	0.25	0	0.75	$0	$35	$70	0	0.5	0.5	DON if middle	$52.50	$52.50
9	5	$0	$20	$70	0.25	0.5	0.25	$0	$20	$35	0	0.5	0.5	DON if high	$27.50	$27.50
8	6	$0	$35	$70	0.25	0	0.75	$0	$35	$70	0	0.5	0.5	DON if middle	$52.50	$52.50
10	6	$0	$35	$70	0.5	0	0.5	$0	$35	$70	0	1	0	DON for any outcome	$35.00	$35.00

Appendix B: Detailed Statistical Results (NOT FOR PUBLICATION)

Table B1: Estimates for Interest Compounding Question

(True response was £110.4)

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
female	.1208637	.1390905	0.87	0.385	-.1517486	.393476
ageS	.3968186	.1936625	2.05	0.040	.017247	.7763901
single	.6368908	.4479348	1.42	0.155	-.2410453	1.514827
white	.2266173	.1577412	1.44	0.151	-.0825497	.5357843
finance	.1057089	.1583877	0.67	0.505	-.2047253	.416143
non_eu	-.0022251	.127354	-0.02	0.986	-.2518344	.2473843
smoker	-.315451	.3116881	-1.01	0.312	-.9263484	.2954465
crtS	-.0427266	.0771227	-0.55	0.580	-.1938843	.1084311
berlinS	.0508408	.0829787	0.61	0.540	-.1117944	.213476
_cons	109.71	.4570172	240.06	0.000	108.8142	110.6057
sigma	.9485433	.1233048			.7352012	1.223793

Table B2: Estimates for Inflation Question

(True response was £198)

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
female	.6555166	.30101	2.18	0.029	.0655479	1.245485
ageS	-.049693	.2060632	-0.24	0.809	-.4535694	.3541834
single	-.0346816	.4901256	-0.07	0.944	-.9953101	.9259469
white	-.1477976	.4230715	-0.35	0.727	-.9770024	.6814072
finance	-.3338726	.2534093	-1.32	0.188	-.8305456	.1628005
non_eu	-.7831535	.4698084	-1.67	0.096	-1.703961	.1376539
smoker	-.095918	.3464372	-0.28	0.782	-.7749224	.5830864
crtS	.0795068	.1481538	0.54	0.592	-.2108692	.3698829
berlinS	.0675041	.1527122	0.44	0.658	-.2318064	.3668145
_cons	199.1655	.601727	330.99	0.000	197.9862	200.3449
sigma	1.696348	.1015747			1.508504	1.907584

Table B3: Estimates for Remaining Lifetime of Men Question

(True response was 59.1 years)

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
female	2.262256	2.497932	0.91	0.365	-2.633602	7.158114
ageS	2.446757	2.312314	1.06	0.290	-2.085295	6.978809
single	-3.488838	7.019735	-0.50	0.619	-17.24727	10.26959
white	7.081467	3.446061	2.05	0.040	.3273123	13.83562
finance	3.842162	2.688428	1.43	0.153	-1.427061	9.111385
non_eu	1.31051	3.622383	0.36	0.718	-5.789229	8.41025
smoker	.396316	2.321885	0.17	0.864	-4.154495	4.947126
crtS	-.9155401	1.338364	-0.68	0.494	-3.538686	1.707606
berlinS	1.790858	1.455642	1.23	0.219	-1.062149	4.643865
_cons	48.28431	7.496065	6.44	0.000	33.5923	62.97633
sigma	15.8356	.8974917			14.17073	17.69607

Table B4: Estimates for Remaining Lifetime of Women Question

(True response was 62.9 years)

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
female	.1826926	2.753928	0.07	0.947	-5.214907	5.580292
ageS	4.212811	1.999146	2.11	0.035	.2945579	8.131064
single	4.533479	5.629901	0.81	0.421	-6.500924	15.56788
white	4.043839	3.208968	1.26	0.208	-2.245623	10.3333
finance	2.796309	2.418511	1.16	0.248	-1.943885	7.536503
non_eu	-.0591597	3.629117	-0.02	0.987	-7.172098	7.053779
smoker	1.303762	2.747118	0.47	0.635	-4.08049	6.688014
crtS	-.3553245	1.254384	-0.28	0.777	-2.813871	2.103222
berlinS	.4161316	1.42578	0.29	0.770	-2.378345	3.210608
_cons	47.63064	7.041492	6.76	0.000	33.82957	61.43171
sigma	15.29233	.7719308			13.8518	16.88265

Table B5: Estimates for Breast Cancer with Probability Format

(True response was 7.8%)

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
female	-2.932982	6.089349	-0.48	0.630	-14.86789	9.001922
ageS	-3.325164	4.301973	-0.77	0.440	-11.75688	5.106547
single	-3.255003	12.89618	-0.25	0.801	-28.53106	22.02105
white	19.23942	8.069034	2.38	0.017	3.424408	35.05444
finance	-13.33804	5.579889	-2.39	0.017	-24.27442	-2.401656
non_eu	14.90143	8.98151	1.66	0.097	-2.702003	32.50487
smoker	21.24842	8.23438	2.58	0.010	5.109333	37.38751
crtS	1.437001	3.130013	0.46	0.646	-4.697711	7.571714
berlinS	-6.916651	3.158586	-2.19	0.029	-13.10737	-.7259355
_cons	27.7921	15.25619	1.82	0.069	-2.109474	57.69367
sigma	31.42903	1.165036			29.22658	33.79745

Table B6: Estimates for Breast Cancer with Frequency Format

(True response was 7.8%)

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
female	1.11684	6.0696	0.18	0.854	-10.77936	13.01304
ageS	-1.122605	4.138559	-0.27	0.786	-9.234031	6.98882
single	-.0654885	11.72954	-0.01	0.996	-23.05497	22.92399
white	8.896363	7.584211	1.17	0.241	-5.968417	23.76114
finance	-4.311176	5.455328	-0.79	0.429	-15.00342	6.38107
non_eu	2.06834	8.455557	0.24	0.807	-14.50425	18.64093
smoker	12.22554	8.624706	1.42	0.156	-4.678573	29.12965
crtS	.0828389	3.335038	0.02	0.980	-6.453715	6.619393
berlinS	-7.323202	3.126175	-2.34	0.019	-13.45039	-1.196011
_cons	24.38284	14.28641	1.71	0.088	-3.618005	52.38369
sigma	30.736	1.488787			27.95226	33.79697

Table B7: Estimates for Breast Cancer with Probability Format

When this Format was the First Question (N=58)

(True response was 7.8%)

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
female	.8342124	9.343188	0.09	0.929	-17.4781	19.14652
ageS	-4.068261	7.706913	-0.53	0.598	-19.17353	11.03701
single	-8.673621	23.61839	-0.37	0.713	-54.96482	37.61758
white	23.82158	12.03242	1.98	0.048	.238464	47.4047
finance	-18.53848	8.124593	-2.28	0.023	-34.46239	-2.614568
non_eu	9.520125	13.93129	0.68	0.494	-17.7847	36.82495
smoker	23.90209	11.11602	2.15	0.032	2.115084	45.6891
crtS	5.531512	4.822146	1.15	0.251	-3.91972	14.98274
berlinS	-7.386577	4.161558	-1.77	0.076	-15.54308	.769927
_cons	34.36851	28.3536	1.21	0.225	-21.20353	89.94055
sigma	30.78217	1.849657			27.36224	34.62954

Table B8: Estimates for Breast Cancer with Frequency Format

When this Format was the First Question (N=62)

(True response was 7.8%)

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
female	5.427551	7.564023	0.72	0.473	-9.397662	20.25276
ageS	-2.164721	5.068092	-0.43	0.669	-12.098	7.768556
single	3.165941	13.94157	0.23	0.820	-24.15903	30.49091
white	16.12129	10.1311	1.59	0.112	-3.735306	35.97788
finance	-5.563114	8.509187	-0.65	0.513	-22.24081	11.11458
non_eu	13.05352	10.7716	1.21	0.226	-8.058435	34.16547
smoker	8.322137	11.58149	0.72	0.472	-14.37716	31.02143
crtS	-5.62591	3.86677	-1.45	0.146	-13.20464	1.95282
berlinS	-4.30888	4.130409	-1.04	0.297	-12.40433	3.786573
_cons	7.952285	15.99824	0.50	0.619	-23.40368	39.30825
sigma	29.5268	2.181793			25.54579	34.12819

Table B9: Estimates for Credit Risk with Probability Format

(True response was 66.7%)

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
female	-12.33833	5.135729	-2.40	0.016	-22.40417	-2.272485
ageS	4.46227	3.943151	1.13	0.258	-3.266165	12.1907
single	10.43192	10.37032	1.01	0.314	-9.893539	30.75737
white	-.5145085	8.351986	-0.06	0.951	-16.8841	15.85508
finance	-6.019441	4.601255	-1.31	0.191	-15.03773	2.998853
non_eu	-3.495325	8.554781	-0.41	0.683	-20.26239	13.27174
smoker	3.003925	6.618789	0.45	0.650	-9.968663	15.97651
crtS	-.0903645	2.920145	-0.03	0.975	-5.813744	5.633015
berlinS	3.56716	2.938142	1.21	0.225	-2.191493	9.325812
_cons	61.81362	13.28393	4.65	0.000	35.77761	87.84964
sigma	26.77588	1.207207			24.51133	29.24965

Table B10: Estimates for Credit Risk with Frequency Format

(True response was 66.7%)

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
female	-7.799498	5.662294	-1.38	0.168	-18.89739	3.298395
ageS	5.214897	3.378843	1.54	0.123	-1.407513	11.83731
single	20.18916	9.327712	2.16	0.030	1.907185	38.47114
white	12.51622	8.802366	1.42	0.155	-4.736099	29.76854
finance	-2.562685	4.959702	-0.52	0.605	-12.28352	7.158153
non_eu	18.62015	9.214859	2.02	0.043	.5593618	36.68095
smoker	.2132735	7.126544	0.03	0.976	-13.7545	14.18104
crtS	.974174	2.796107	0.35	0.728	-4.506094	6.454442
berlinS	5.769675	2.692328	2.14	0.032	.4928082	11.04654
_cons	28.06668	13.17115	2.13	0.033	2.251706	53.88165
sigma	27.13271	1.296473			24.70703	29.79654

Table B11: Estimates for Credit Risk with Probability Format

When this Format was the First Question (N=62)

(True response was 66.7%)

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf. Interval]	
female	-10.12084	7.174753	-1.41	0.158	-24.1831	3.941413
ageS	8.998826	3.891018	2.31	0.021	1.372572	16.62508
single	26.74031	9.910682	2.70	0.007	7.315727	46.16489
white	10.23155	8.743735	1.17	0.242	-6.905859	27.36895
finance	-7.732764	6.944066	-1.11	0.265	-21.34288	5.877355
non_eu	2.997623	9.175632	0.33	0.744	-14.98629	20.98153
smoker	4.390203	6.917286	0.63	0.526	-9.167428	17.94783
crtS	-4.006747	3.657293	-1.10	0.273	-11.17491	3.161416
berlinS	11.29258	2.940537	3.84	0.000	5.52923	17.05592
_cons	34.54933	14.27547	2.42	0.016	6.569911	62.52874
sigma	25.57656	1.413615			22.95072	28.50283

Table B12: Estimates for Credit Risk with Frequency Format

When this Format was the First Question (N=62)

(True response was 66.7%)

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf. Interval]	
female	-1.50905	8.081353	-0.19	0.852	-17.34821	14.33011
ageS	9.666001	4.593329	2.10	0.035	.6632411	18.66876
single	22.37095	15.45922	1.45	0.148	-7.928573	52.67047
white	8.429374	16.17703	0.52	0.602	-23.27702	40.13577
finance	-1.227179	7.491029	-0.16	0.870	-15.90933	13.45497
non_eu	7.642467	17.16264	0.45	0.656	-25.99569	41.28062
smoker	-.909114	14.00157	-0.06	0.948	-28.35168	26.53345
crtS	2.046534	4.300484	0.48	0.634	-6.382259	10.47533
berlinS	4.751404	4.278632	1.11	0.267	-3.634561	13.13737
_cons	29.81129	21.9804	1.36	0.175	-13.26949	72.89208
sigma	27.72974	1.81675			24.38812	31.52923

Table B13: Estimates for Breast Cancer Risk with Both Formats

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
frequency	-8.780653	2.622727	-3.35	0.001	-13.9211	-3.640202
breast_F_first	-1.880165	5.001399	-0.38	0.707	-11.68273	7.922396
female	-.8176635	5.257625	-0.16	0.876	-11.12242	9.487092
ageS	-1.958312	3.738468	-0.52	0.600	-9.285575	5.368951
single	-1.349017	10.54452	-0.13	0.898	-22.0159	19.31787
white	14.54114	5.662997	2.57	0.010	3.441867	25.64041
finance	-8.783428	4.979758	-1.76	0.078	-18.54357	.9767177
non_eu	8.731927	6.837669	1.28	0.202	-4.669657	22.13351
smoker	16.56504	7.027368	2.36	0.018	2.791648	30.33843
crtS	.7693924	2.857754	0.27	0.788	-4.831702	6.370486
berlinS	-7.244087	2.870351	-2.52	0.012	-12.86987	-1.618302
_cons	30.79826	12.71119	2.42	0.015	5.884786	55.71173
sigma	31.26975	1.176413			29.04698	33.66262

Table B14: Estimates for Breast Cancer Risk with Both Formats

When the Frequency Format was the Second Question (N=58)

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
frequency	-7.906759	3.865179	-2.05	0.041	-15.48237	-.331147
female	-.5126137	8.652634	-0.06	0.953	-17.47146	16.44624
ageS	-.0453931	6.898456	-0.01	0.995	-13.56612	13.47533
single	-1.241718	21.8429	-0.06	0.955	-44.05301	41.56957
white	13.45107	10.82962	1.24	0.214	-7.774593	34.67674
finance	-12.97861	6.627994	-1.96	0.050	-25.96924	.0120169
non_eu	1.854472	11.93734	0.16	0.877	-21.54228	25.25122
smoker	16.85493	10.75168	1.57	0.117	-4.217971	37.92782
crtS	6.294715	4.243566	1.48	0.138	-2.022523	14.61195
berlinS	-8.572086	3.629374	-2.36	0.018	-15.68553	-1.458644
_cons	37.24837	26.07599	1.43	0.153	-13.85964	88.35638
sigma	31.09647	1.710031			27.91917	34.63536

Table B15: Estimates for Credit Risk with Both Formats

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
frequency	-1.093302	2.457748	-0.44	0.656	-5.910399	3.723795
credit_F_first	3.262956	4.167427	0.78	0.434	-4.90505	11.43096
female	-9.885857	4.318312	-2.29	0.022	-18.34959	-1.42212
ageS	4.543536	2.863814	1.59	0.113	-1.069436	10.15651
single	15.09596	6.717241	2.25	0.025	1.930411	28.26151
white	6.205505	7.602942	0.82	0.414	-8.695988	21.107
finance	-4.698919	4.093765	-1.15	0.251	-12.72255	3.324713
non_eu	7.952487	7.896865	1.01	0.314	-7.525083	23.43006
smoker	2.008373	5.896095	0.34	0.733	-9.547762	13.56451
crtS	.2163768	2.468343	0.09	0.930	-4.621486	5.054239
berlinS	4.603194	2.10187	2.19	0.029	.4836052	8.722783
_cons	43.59671	11.69502	3.73	0.000	20.67488	66.51854
sigma	27.23843	1.107556			25.15191	29.49805

Table B16: Estimates for Credit Risk with Both Formats

When the Frequency Format was the Second Question (N=62)

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
frequency	3.51211	3.540208	0.99	0.321	-3.426569	10.45079
female	-11.10497	5.873301	-1.89	0.059	-22.61643	.4064855
ageS	4.224511	3.454046	1.22	0.221	-2.545295	10.99432
single	22.86012	6.34174	3.60	0.000	10.43053	35.2897
white	15.25103	8.144386	1.87	0.061	-.7116762	31.21373
finance	-5.099743	5.762526	-0.88	0.376	-16.39409	6.1946
non_eu	18.05947	9.300997	1.94	0.052	-.1701477	36.28909
smoker	2.293681	5.355934	0.43	0.668	-8.203757	12.79112
crtS	-1.434976	3.189159	-0.45	0.653	-7.685613	4.81566
berlinS	8.230033	2.312235	3.56	0.000	3.698136	12.76193
_cons	25.1368	12.61047	1.99	0.046	.420728	49.85287
sigma	26.2472	1.390422			23.65872	29.11888

Figure B1: Distribution of RRA
Based on Individual Model of Behavior

Estimates obtained for 114 out of 120 subjects.